Mathematics
QA
1
A461
MAY 1 1937
C-2

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

E. T. BELL
CALIFORNIA INSTITUTE OF TECHNOLOGY

ABRAHAM COHEN
THE JOHNS HOPKINS UNIVERSITY

E. W. CHITTENDEN
UNIVERSITY OF IOWA

F. D. MURNAGHAN
THE JOHNS HOPKINS UNIVERSITY

J. F. RITT
COLUMBIA UNIVERSITY

WITH THE COÖPERATION OF

FRANK MORLEY
W. A. MANNING
HARRY BATEMAN
HARRY LEVY
J. R. KLINE

MARSTON MORSE
E. P. LANE
ALONZO CHURCH
L. R. FORD
OSCAR ZARISKI

G. C. EVANS
AUREL WINTNER
GABRIEL SZEGŐ
R. L. WILDER
R. D. JAMES

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LIX, Number 2
APRIL, 1937

THE JOHNS HOPKINS PRESS
BALTIMORE, MARYLAND
U. S. A.

CONTENTS

THE AMERICAN JOURNAL OF MATHEMATICS appears four times yearly.

The subscription price of the JOURNAL for the current volume is \$7.50 (foreign postage 50 cents); single numbers \$2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications may be sent to Dr. A. COHEN at The Johns Hopkins University.

Subscriptions for the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice. acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

GE

35
35

1

0

5
0

5

6

5
7

ASTRONOMICAL CONSEQUENCES OF THE RELATIVISTIC TWO-BODY PROBLEM.*

By Tullio Levi-Civita.

1. Mechanical laws, according to Einstein's theory, are much more complicated in conception than under the assumptions of Newton. However the motion of celestial bodies under ordinary circumstances differ so little from their Newtonian representation, that, for astronomical purposes, relativistic effects may be conveniently treated as first-order perturbations.

A good amount of work in this direction was done, shortly after the appearance of general relativity, with deep insight and high competence by the late Professor De Sitter.[1]

The simple case of two bodies of *comparable* masses lies beyond De Sitter's developements, which were chiefly directed towards the inclusion of perturbations arising from relativity in the standard equations concerning planets and satellites of our solar system, where one of the masses predominates.

I have recently taken up the question,[2] paying due attention to the case of comparable masses. For the usual two-body problem, which in the traditional hierarchy comes immediately after Einstein's one-centre problem, the equations of motion are certainly integrable if one treats relativistic effects as first-order perturbations.

We intend to say a few words about deduction and illustration of two inequalities which are already apparent or, at least, may shortly appear in the observable field.

2. Let us start from the explicit form of the two Lagrangian functions

$$L_0 \text{ and } L_1$$

which define the *absolute* motion of the centres of mass, P_0 and P_1, of two celestial bodies; e. g., a double star.

I suppose that everything has already been reduced to ordinary space, $x_h{}^i$

* A paper delivered at the Tercentenary Conference of Arts and Sciences at Harvard University, September 4, 1936. Received by the Editors January 18, 1937.

[1] *Monthly Notices, Royal Astronomical Society,* vol. 77 (1916), pp. 155-184 (Second paper).

[2] "The relativistic problem of several bodies," *American Journal of Mathematics,* vol. 59 (1937), pp. 9-22.

($i = 1, 2, 3$) being Cartesian coördinates of P_h ($h = 0, 1$) with reference to some fixed or Galilean frame, while the independent variable is $x^0 = ct$ (t usual time and c velocity of light).

The Lagrangian equations, furnished by the L_h ($h = 0, 1$), define the components

$$\ddot{x}_h{}^i \qquad (h = 0, 1)$$

of the absolute accelerations of the two points P_h as functions of positions and velocities. At this stage, all textbooks introduce relative coördinates

$$x^i = x_1{}^i - x_0{}^i \qquad (i = 1, 2, 3)$$

and corresponding relative accelerations

$$\ddot{x}^i = \ddot{x}_1{}^i - \ddot{x}_0{}^i \qquad (i = 1, 2, 3)$$

simply by subtraction. Then, on account of the fact that all but the Newtonian terms are of the second order, we are allowed to use Keplerian values, and especially to employ the classical integrals of energy and areas.

If the simple but rather tedious developments are performed with some insight into the matter, we recognize that relative motion may also be brought under the Lagrangian scheme. The corresponding function L is

$$(1) \qquad L = N + \Pi,$$

where, to within a constant factor, N designates the usual Newtonian term and Π the additional relativistic contribution. More precisely, if m_0, m_1 are the masses of the two bodies and $r = \overline{P_0P_1}$ is their mutual distance, then putting

$$(2) \qquad m = m_0 + m_1$$

and

$$(3) \qquad \beta_i = \frac{dx^i}{dx^0} = \dot{x}^i, \qquad \beta^2 = \sum_1^3{}_i \beta_i{}^2, \qquad \gamma = \frac{fm}{c^2}\,\frac{1}{r},$$

we obviously have

$$(4) \qquad N = \tfrac{1}{2}\beta^2 + \gamma.$$

In the expression of Π we shall denote by $\mathfrak{e}$ the difference

$$\tfrac{1}{2}\beta^2 - \gamma$$

which, in Newtonian approximation, is nothing but the constant of energy divided by c^2, so that, up to terms of second order, the numerical value of $\mathfrak{e}$ behaves like a constant. On the other hand, it is not the same to apply the Lagrangian operator

$$\frac{d}{dx^0}\frac{\partial}{\partial\beta_i}-\frac{\partial}{\partial x^i}$$

to the constant $\mathfrak{e}$, which gives zero, as to the binomial $\frac{1}{2}\beta^2-\gamma$, which gives $\dot{\beta}_i+\partial\gamma/\partial x^i$, or, up to the first order, $2\partial\gamma/\partial x^i$. Therefore, if we include $\mathfrak{e}$ among the arguments by means of which the explicit expression of Π is built, it is necessary to state whether $\mathfrak{e}$ is to be treated, in performing Lagrangian operations, as a genuine constant or as the difference $\frac{1}{2}\beta^2-\gamma$, which, as far as first approximation is concerned, has the same numerical value.

Attributing to $\mathfrak{e}$ at any moment the rôle of a simple constant, I have obtained

$$\Pi=(2-\tfrac{1}{2}\mathfrak{p})\beta^2\gamma-(1+\tfrac{1}{2}\mathfrak{p})\gamma^2+2(-1+2\mathfrak{p})\mathfrak{e}\gamma+\tfrac{1}{2}\mathfrak{p}\frac{1}{\mathfrak{a}^2}\gamma^3, \tag{5}$$

where

$$\mathfrak{p}=\frac{m_0m_1}{m^2} \tag{6}$$

and $1/\mathfrak{a}$ is a further dimensionless constant, connected, within terms of higher order, with the double areal velocity C by the relation

$$\frac{1}{\mathfrak{a}}=\frac{C}{c.fm/c^2}\,. \tag{7}$$

Accordingly, in ordinary planetary pairs and double stars, $\mathfrak{a}\sim\beta$, i.e., $\mathfrak{a}$ has the same order of magnitude as β. In particular, for circular motion, $\mathfrak{a}$ is the constant value of β. This is verified at once by means of the elementary relations

$$\beta^2=\frac{fm}{c^2}\frac{1}{a},\qquad \frac{C}{c}=a\beta,$$

in which a means the radius or, more generally, the major semi-axis of the orbit.

Solving for fm/c^2 and C/c, and substituting in (7), we get simply

$$\frac{1}{\mathfrak{a}}=\frac{1}{\beta}\,.$$

At any rate, from $\mathfrak{a}\sim\beta$ and $\gamma\sim\beta^2$, it follows that

$$\frac{1}{\mathfrak{a}^2}\gamma^3\sim\gamma^2,$$

showing that the last term of Π is, like the others, of the second order.

Squaring (7) and remembering the classical relation

$$C^2 = fma(1 - e^2) \qquad (e = \text{eccentricity}),$$

we get

$$\frac{1}{\mathfrak{a}^2}\,\frac{fm}{c^2} = a(1 - e^2), \tag{7'}$$

which will be used later on.

3. The expression (1) of L has, in view of (3) and (5), the form

$$L = \tfrac{1}{2}\psi\beta^2 + \phi, \tag{8}$$

where both ψ and ϕ depend exclusively on the mutual distance r, since

$$\psi = 1 + (4 - \mathfrak{p})\gamma, \tag{9}$$

$$\phi = \gamma - (1 + \tfrac{1}{2}\mathfrak{p})\gamma^2 + 2(-1 + 2\mathfrak{p})\mathfrak{e}\gamma + \tfrac{1}{2}\mathfrak{p}\frac{1}{\mathfrak{a}^2}\gamma^3. \tag{10}$$

The motion defined by the Lagrangian function L admits the integral

$$\tfrac{1}{2}\psi\beta^2 - \phi = \mathfrak{e}^*, \tag{11}$$

where the constant $\mathfrak{e}^*$ may differ from $\mathfrak{e}$ only in second order terms.

Now an equivalence theorem in analytical dynamics [3] states that the Lagrangian function

$$L_1 = \tfrac{1}{2}\beta^2 + \psi(\phi + \mathfrak{e}^*), \tag{12}$$

for which the motion admits the integral

$$\tfrac{1}{2}\beta^2 - \psi(\phi + \mathfrak{e}^*) = \text{const.}, \tag{13}$$

gives rise, for the value 0 of the constant in the second member, to a family of trajectories identical with those defined by (8) and the integral (11).

Therefore, as far as trajectories are concerned, our task is reduced to characterize those belonging to the Lagrangian function (12). The latter corresponds to the motion of a free particle in ordinary space under a conservative (even central) force having the force-function

$$\Phi = \psi(\phi + \mathfrak{e}^*).$$

Omitting the additive constant $\mathfrak{e}^*$ and all terms of order higher than two, and writing correspondingly $\mathfrak{e}$ instead of $\mathfrak{e}^*$ in all second order terms, we have, by (9) and (10),

---

[3] Cf., e. g., Levi-Civita and Amaldi, *Lezioni di meccanica razionale*, vol. II_2 (Bologna, Zanichelli, 1927), pp. 514-515.

$$\Phi = \gamma + (2 + 3\mathfrak{p})\mathfrak{e}\gamma + 3(1 - \tfrac{1}{2}\mathfrak{p})\gamma^2 + \frac{\mathfrak{p}}{\mathfrak{a}^2}\gamma^3.$$

Putting for a moment

$$m^* = m\{1 + (2 + 3\mathfrak{p})\mathfrak{e}\} \quad \text{and} \quad \gamma^* = \frac{fm^*}{c^2}\,\frac{1}{r} = \gamma + (2 + 3\mathfrak{p})\mathfrak{e}\gamma,$$

we may write simply γ^* instead of γ in the second and third terms of Φ. Here, however, it is indifferent, up to the second order, whether one employs γ or γ^*. Hence, omitting asterisks, we may consider the trajectories of a central force with the potential function

$$\text{(I)} \qquad \gamma + 3(1 - \tfrac{1}{2}\mathfrak{p})\gamma^2 + \tfrac{1}{2}\frac{\mathfrak{p}}{\mathfrak{a}^2}\gamma^3,$$

where, from (2), (3), (6) and (7),

$$\text{(II)} \quad m = m_0 + m_1, \qquad \gamma = \frac{fm}{c^2}\,\frac{1}{r}, \qquad \mathfrak{p} = \frac{m_0 m_1}{m^2}, \qquad \frac{1}{\mathfrak{a}} = \frac{C}{cfm/c^2}.$$

Note that, in these formulae (owing to the described policy of first introducing m^* and γ^* and then suppressing asterisks), m_0 and m_1 do not represent exactly the ordinary masses of our two bodies (which had been introduced in **2**), but truly these masses, slightly altered by the constant factor

$$1 + (2 + 3\mathfrak{p})\mathfrak{e}.$$

What essentially matters is that they, like their sum m, behave as constants driving the motions to be now considered.

Obviously, the first term in (I) represents the Newtonian attraction, while the other two (both of the second order) are the relativistic perturbations consisting of central attractions; the one varying according the inverse cube, the other according the inverse fourth power of the distance. For the Einsteinian case of one-centre problem, one has only to put $\mathfrak{p} = 0$, and (I) gives the well known expression $3\gamma^2$ for the perturbative function.

4. Orbits described under central forces were thoroughly investigated in the 18th and 19th centuries. Especially for orbits which may be regarded as disturbed Keplerian ellipses, computations of apsidal angles, and corresponding precessions of perihelia may be obtained by elementary methods. In this way, with first order accuracy, the angular precession (per revolution) of the perihelion or, in the case of a double star, of the periastron, is found to be

$$\text{(III)} \qquad \sigma = \sigma_e = 6\pi\mathfrak{a}^2.$$

With the value (7′) of $\mathfrak{a}^2$, namely

$$\frac{fm}{c^2}\,\frac{1}{a(1-e^2)},$$

the expression $\sigma_e = 6\pi\mathfrak{a}^2$ *of* σ *is exactly the precession predicted by Einstein for an infinitesimal planet* ($\mathfrak{p} = 0$) *in the case of motion about a central mass possessing the total mass* m *of the binary system.* Therefore, within the required approximation, Einstein's formula, first established for an infinitesimal body in the (relativistic) field of a central mass m, is still valid for two bodies of any masses m_0, m_1 whose sum is m.

I hope that, for some double stars, the precession of the periastron of the satellite star may be observed with sufficient accuracy to test the theoretical result, thus affording a new astronomical confirmation of Einstein's gravitational theory. For the moment I can only draw attention to the matter.

5. The above prediction refers to *relative* orbits. Another, and perhaps more striking, theoretical deduction concerns the *absolute* motion in the sky of any double star system.

It is well known that general relativity does not include, as a rigorous law, the principle of reaction nor its most popular dynamical consequence, concerning the motion of the center of mass in case of absence of external forces. Accordingly, we can no longer rely upon the constance of the absolute velocity $\dot{G} = dG/dx^0$ of the centre of mass G of a double star, but are, on the other hand, enabled to infer the expression of its instantaneous (absolute) acceleration by employing the relativistic treatment of the two-body problem.

The general idea is obvious. Having the Lagrangian equations of motion for the two bodies P_0 and P_1, equations obtained in my previous paper, we deduce at once the vectors $\dot{\boldsymbol{\beta}}_0$ and $\dot{\boldsymbol{\beta}}_1$ of the absolute accelerations as functions of their *relative* positions, and (still) absolute velocities.

The combination

$$\alpha = \frac{1}{m}\,(m_0\boldsymbol{\beta}_0 + m_1\boldsymbol{\beta}_1)$$

is precisely the velocity of G when referred to x^0 as the time variable, i. e., the ordinary velocity divided by c; while the acceleration of G, again referred to x^0, is

$$\dot{\alpha} = \frac{1}{m}\,(m_0\dot{\boldsymbol{\beta}}_0 + m_1\dot{\boldsymbol{\beta}}_1)\,, \tag{14}$$

where the dot denotes d/dx^0.

In view of the classical mechanics, which always holds in the first approximation, we may anticipate that the Newtonian terms in the right-hand member of (14) disappear, so that there remains only the relativistic correc-

tion, expressed, as before, in terms of relative positions and absolute velocities. Now, if we consider actual double star motions, it is plainly permitted, within the degree of approximation in which we are interested, to introduce Newtonian values referring to Keplerian motions of negative energy.

In order to perform the computation along the lines indicated above, it will be convenient to use relative coördinates of invariable direction and having their origin in P_0, where m_0 is the *principal star* ($m_0 \geq m_1$), and to choose the orthogonal trihedron $P_0x^1x^2x^3$ in its standard position: P_0x^1 towards the periastron of the (undisturbed) elliptical orbit of P_1; P_0x^2 in the plane of this orbit and rotated 90° in the sense of the motion; P_0x^3 forming a right-handed trihedron with the preceding two.

First we recognize from the outlined formulae that *the component* $\dot{\alpha}_3 = 0$. *Therefore, the acceleration of the center of mass G of a double star lies entirely in the plane of* (relative) *orbit*; we may also say that it lies in the common plane of (absolute) *orbits*, described by P_0 and P_1 about G.

For the two *components*

$$\dot{\alpha}_1 = \frac{d\alpha_1}{dx^0}, \qquad \dot{\alpha}_2 = \frac{d\alpha_2}{dx^0}$$

in the orbital plane I have found

$$(15) \qquad \frac{d\alpha_i}{dx^0} = \mathfrak{p}\mathfrak{d}\left\{-\frac{d}{dx^0}(\gamma\beta_i) + \frac{\partial\gamma}{\partial x^i}\left(\mathfrak{e} + 4\gamma - \tfrac{3}{2}\frac{1}{\mathfrak{a}^2}\gamma^2\right)\right\}, \quad (i = 1, 2),$$

where the factor

$$(16) \qquad \mathfrak{d} = \frac{m_0 - m_1}{m}$$

is proportional to the difference of the masses, while β_i and γ have the same significance as in (3), $\mathfrak{p}$ being defined by (6), $\mathfrak{a}$ by (7), while $\mathfrak{e}$ is the (negative) total energy of the undisturbed Keplerian motion.

6. As already remarked, Keplerian values when used in the right-hand members of (15) give sufficient accuracy. Then the explicit determination of the two variable components of velocity, $\alpha_i(x^0)$, requires only quadratures, easily performed by introducing the true anomaly θ instead of x^0 by means of the relation

$$r^2\frac{d\theta}{dx^0} = \frac{C}{c} = \frac{\sqrt{fma(1-e^2)}}{c}.$$

and remembering that

$$\gamma = \frac{fm}{c^2}\frac{1}{r} = \frac{fm}{c^2}\frac{1 + e\cos\theta}{a(1-e^2)}.$$

Periodical terms in θ correspond to small fluctuations in the components α_1 and α_2, fluctuations which are repeated during every revolution and certainly remain within the limits of accuracy of observation in the case of all the known double stars. Accordingly, the only interesting terms are the *secular terms,* whose effects accumulate during the successive revolutions. Now we must remember that α_i mean components of velocity with respect to the (Roemerian) time $x^0 = ct$. Therefore *the components of the ordinary velocity of G are $c\alpha_i$.*

Denoting by $c\bar{\alpha}_1$ and $c\bar{\alpha}_2$ their secular parts in terms of θ, we obtain finally, in virtue of (7′),

$$c\bar{\alpha}_1 = -\tfrac{1}{2}\mathfrak{p}\mathfrak{d}\frac{e}{(1-e^2)^{3/2}}\frac{fm}{c^2}\sqrt{\frac{fm}{a^3}}\,\theta, \qquad c\bar{\alpha}_2 = 0.$$

The final conclusion is that *the secular acceleration of the center of mass G of the double star is directed along the major axis towards the periastron of the principal star.* The amount of this secular acceleration may be conveniently expressed as the increase of velocity in Km/sec per revolution.

To this end, we first introduce the mass of the Sun, $m_\odot$, and write

$$\frac{fm}{c^2} = \frac{m}{m_\odot}\,\frac{fm_\odot}{c^2}.$$

The second factor is a length, the so called *gravitational radius* $l_\odot$ of the Sun, having a value of the order of magnitude of a Kilometer, or about 1, 5 Km. We may then write

$$\frac{fm}{c^2} = \frac{m}{m_\odot}\cdot 1,5 \text{ Km}.$$

On the other hand, the mean motion $\sqrt{\frac{fm}{a^3}}$ of the double star is $\frac{2\pi}{T}$, where T is the period of revolution. Of course, T refers to the unit of time used previously in f and c. Starting with the C. G. S. units, we have T in seconds. But, in data of double stars, T is generally expressed in days (for spectroscopic binaries) or in years (for visual binaries). Choosing the first case and writing T^d to avoid ambiguity, we have

$$\frac{1}{T} = \frac{1}{T^d}\,\frac{1}{86164}.$$

It follows then, putting $\theta = 2\pi$ in the preceding expression of $c\bar{\alpha}_1$ and considering its absolute value V, that the increase ΔV of the velocity of G during a revolution is

$$\text{(IV)} \qquad (\Delta V)_{\text{per revolution}} = \tfrac{1}{2}\mathfrak{p}\mathfrak{d}\,\frac{e}{(1-e^2)^{3/2}}\,\frac{m}{m_\odot}\,\frac{4\pi^2}{86164}\,\frac{1,5}{T^d}\,\text{Km/sec.}$$

The number of revolutions per day is $1/T^d$, and, per century, $100 \cdot 365,25$ times $1/T^d$. Since

$$\tfrac{1}{2}\cdot\frac{4\pi^2}{86164}\cdot 1,5 \cdot 100 \cdot 365,25 = 12,55,$$

the increase of the velocity of the center of mass during a century is

$$\text{(V)} \qquad (\Delta V)_{\text{in a century}} = 12,55\mathfrak{p}\mathfrak{d}\,\frac{e}{(1-e^2)^{3/2}}\,\frac{m}{m_\odot}\,\frac{1}{(T^d)^2}\,\text{Km/sec.}$$

7. Such a difference of velocity along the apsidal line, having a component also in the line of sight, ought to be detectable eventually by spectroscopic observation.

As far as numerical values are concerned, we recognize from (V) [or from (IV)] that the most favorable circumstances are realized for double stars having the following properties:

a) short period, i. e., stars very near each other, which strongly influences $1/T^d$;

b) total mass m, large (or at least not too small) in comparison with the mass of the Sun $m_\odot$;

c) pronounced eccentricity, on account of the factor $e/(1-e^2)^{3/2}$;

d) comparable masses, but not nearly equal, owing to the factors

$$\mathfrak{p} = \frac{m_0 m_1}{(m_0+m_1)^2}, \qquad \mathfrak{d} = \frac{m_0 - m_1}{m_0 + m_1}.$$

The best conditions for $\mathfrak{p}\mathfrak{d}$ are realized by mass-ratios

$$\frac{m_1}{m_0+m_1} = x, \qquad \frac{m_0}{m_0+m_1} = 1-x$$

for which the polynomial

$$\mathfrak{p}\mathfrak{d} = x(1-x)(1-2x)$$

attains its maximum. This takes places for $x = \frac{1}{2}(1-3^{-1/2})$, or, roughly, for two stars containing respectively ¼ and ¾ of the total mass of the system. The corresponding value of $\mathfrak{p}\mathfrak{d}$ is about 0,1.

I am well aware of the astronomical observations which have shown that, in general, the eccentricity e decreases with T^d; so it will not be easy to find a

binary for which the two requirements a) and c) are equally well satisfied. On the other hand, a) predominates in (V), $1/T^d$ appearing to the second power; furthermore, b) is, in the main, in accordance with a).

As the visual binaries have, in general, long periods (some years), they are not to be expected, on account of a), to be advantageous for testing the formula (V). This formula requires, however, the knowledge of the masses of the principal and the companion star. Accordingly, it will be advisable to turn to the class of binaries for which photometric as well as spectrographic observations are available. In order to consider, at least, one example with a reasonable $(\Delta V)_{\text{in a century}}$, I have looked into Moore's Tables of the Lick Observatory,[4] stopping at N° 28, b^1 Persei, for which, unfortunately, only spectroscopic data are certain.

The tabulated elements are

$$m_0 = \frac{0,85}{\sin^3 i} m_\odot, \qquad m_1 = \frac{0,23}{\sin^3 i} m_\odot, \qquad T^d = 1,52, \qquad e = 0,22,$$

i being the *hitherto unknown* inclination of the orbital plane to the tangential plane of the celestial sphere. Whatever i may be, we have finally for b^1 Persei

$$\frac{m}{m_\odot} = \frac{1,08}{\sin^3 i},$$

while $\mathfrak{p}$ and $\mathfrak{d}$, involving only mass-ratio, are independent of i. Their product has the value

$$\mathfrak{p}\mathfrak{d} = 0,09622.$$

Accordingly, formula (V) gives for b^1 Persei

$$(\Delta V)_{\text{in a century}} = 12,55 \cdot 0,09622 \cdot \frac{0,22}{(0,9516)^{3/2}} \frac{1,08}{\sin^3 i} \frac{1}{(1,52)^2} \text{ Km/sec}$$
$$= \frac{0,13}{\sin^3 i} \text{ Km/sec.}$$

The presence of the (yet unknown) divisor $\sin^3 i$ is consistent with the hope that ΔV may become appreciable much earlier than in a century; perhaps even in a few years.

UNIVERSITY OF ROME.

---

[4] Or rather to the Resumé, reported in Armellini's *Astronomia siderale*, vol. II (Bologna, Zanichelli, 1931), Appendix 3.

FINITE DEFORMATIONS OF AN ELASTIC SOLID.*

By F. D. MURNAGHAN.

Introduction. In the classical theory of elasticity a deformation (= strain) is termed infinitesimal when the space derivatives of the components of the displacement vector of an arbitrary particle of the medium are so small that their squares and products may be neglected. Many attempts have been made to extend the classical theory of infinitesimal strain to the case of finite strains i. e. strains in which the fundamental hypothesis which serves to define an infinitesimal strain is not legitimate. The more important of these are given in the references numbered **1** to **7** at the end of the present paper. A good summary with many references may be found in the address of Professor Signorini (**8**) at the Palermo meeting (1935) of the *Societa Italiana per il progresso delle scienze.* In the case of a finite strain there are two essentially different viewpoints which coalesce when the strain is infinitesimal: we may use as the independent variables in terms of which the strain is described either

(a) the coördinates of a typical particle of the medium in the initial or unstrained position or

(b) the coördinates of a typical particle in the final or strained position. Adopting the terminology familiar in the corresponding situation in hydrodynamics we refer to these as the Lagrangian and Eulerian viewpoints respectively. Most of the previous writers on the subject of finite strain have, probably for reasons of mathematical convenience, adopted the Lagrangian view-point but in the present paper (which is concerned with actual applications of the theory) the Eulerian point of view is regarded as fundamentally more significant than the Lagrangian. In this connection the following quotation from the recent paper by Seth (**7**) is to the point:

> "Like the body-stress equations these (the strain components) should be referred to the actual position of a point P of the material in the strained condition, and not to the position of a point considered before strain. The importance of this point, overlooked by various authors, can not be exaggerated. Apparently Filon and Coker (**9**) were the first to notice it and to stress its importance."

In the classical theory one of the fundamental results (derived from the principle of energy conservation) expresses the connection between stress and strain as follows:

* Received February 23, 1937.

The stress tensor equals the gradient of the elastic-energy-density with respect to the strain tensor.

This fundamental principle (which is the formulation of Hooke's law in its most general form) merely states that, in a virtual displacement of the strained elastic medium, the virtual work of all the forces, both surface and body, acting upon the medium may be obtained by integrating over the medium the scalar product of the stress tensor by the variation of the strain tensor. We show in the present paper that this principle is merely an approximation which, whilst valid in the infinitesimal theory, is not valid in the finite theory. The exact principle is that the virtual work is obtained by integrating over the medium the scalar product of the stress tensor by the *space-derivative of the virtual displacement vector* and it is only in the infinitesimal theory that one may with propriety equate the variation of the strain tensor to the space derivative of the virtual displacement vector. It is a fortunate circumstance, the demonstration of which is the raison d'être of the present paper, that the exact equations, valid for any deformation, are sufficiently simple, at least in the case of an isotropic solid, to be applied and to be compared with experimental results. We apply them to Bridgman's experiments with solids and liquids under high pressures (up to 20,000 atmospheres) and find remarkable agreement without introducing more than the two elastic constants of the infinitesimal theory. We also treat the Young's modulus experiment and obtain at least a qualitative explanation of the yield point phenomenon which is not cared for in the classical theory. The mathematical treatment proceeds most naturally and simply when one uses the methods of tensor analysis. It is, however, not necessary to be especially familiar with these methods in order to understand the reasoning and we shall indicate at appropriate places how a non-tensor argument (called, for brevity, Cartesian) would proceed. In the interest of clarity (and probably also of brevity) it has seemed better to make the paper self-contained.

1. The strain tensor and its variation. We are concerned with a three-dimensional medium (which we shall regard as a collection of particles) and with two positions of this medium to which we shall refer as the initial or unstrained and the final or strained position. A typical particle of the medium will have initial and final coördinates. In the classical theory it has been usual to employ the same reference frame (rectangular Cartesian) for both the initial (unstrained) and final (strained) positions and to denote the initial coördinates by (a, b, c) and the final coördinates (of the *same* particle) by (x, y, z). For the treatment we propose here it is inconvenient to tie our hands

at the very beginning by the restrictive hypothesis that the same coördinate reference frame will be used to describe both positions of the medium; the advantage of not making this hypothesis being that it is then possible to make, or contemplate making, transformations of the final coördinates without thereby enforcing a change of the initial coördinates. In the language of tensor analysis the initial coördinates will be invariants or scalars under transformations of the final coördinates. In order to emphasize this fact in the symbolism we shall methodically write the labels necessary to distinguish from one another the various members of a set of scalar quantities to the *left* of the letter which is the symbol for the set; reserving, as is usual, the *right* of a letter for the labels which are necessary to distinguish from one another the various components of a tensor of which the letter is the symbol. Thus we shall denote the initial coördinates of a typical particle of the medium by ${}^r a$ and the final coördinates of the same particle by x^s (the labels r and s running independently over the range 1, 2, 3). The initial coördinates ${}^r a$, as well as the final coördinates x^s, are, independently of each other, any sets of coördinates which we may find convenient; e. g. the initial coördinates may be rectangular Cartesian and the final coördinates space polar. We make the usual assumption that either coördinate system is differentiable (with continuous first derivatives) with respect to the other and we adopt the notations

$$ {}^r a_{,s} = \partial\, {}^r a / \partial x^s ; \qquad {}_{s,} x^r = \partial x^r / \partial\, {}^s a $$

for the first order partial derivatives. As the notation implies ${}^r a_{,s}$ is, for fixed r and varying s, a covariant vector, namely the gradient of the scalar function ${}^r a$ whilst ${}_{s,} x^r$ is, for fixed s and varying r, a contravariant vector (which furnishes the direction of the coördinate line along which ${}^s a$ varies, the other two a's being held constant). The fundamental reciprocal nature of these two vectors is described by the formulae

$$ ({}^r a_{,\sigma})({}_{s,} x^\sigma) = {}^r{}_s \delta ; \qquad ({}_{\sigma,} x^r)({}^\sigma a_{,s}) = \delta_s{}^r . $$

In these formulae we follow the standard convention of tensor analysis according to which a repeated label (in this paper always taken from the Greek alphabet) occurring once above and once below indicates summation with respect to that label over the range 1, 2, 3; and ${}^r{}_s\delta$, $\delta_s{}^r$ each have the value unity when $r = s$ and the value zero otherwise. As the notation implies ${}^r{}_s\delta$ is a set of 9 scalar functions whilst $\delta_s{}^r$ are the 9 components of a single mixed tensor.

The initial and final squared elements of arc length will be given by formulae of the type:

$$ (ds_0)^2 = {}_{\alpha\beta} c \,(d\, {}^\alpha a)(d\, {}^\beta a) ; \qquad (ds)^2 = g_{\alpha\beta}\, dx^\alpha\, dx^\beta $$

which are induced by the postulated underlying Euclidean metric of the space in which our medium is being deformed. For example if both coördinate systems are rectangular Cartesian

$$(ds_0)^2 = (da)^2 + (db)^2 + (dc)^2; \qquad (ds)^2 = (dx)^2 + (dy)^2 + (dz)^2$$

whilst if the initial coördinate system is rectangular Cartesian and the final space polar

$$(ds_0)^2 = (da)^2 + (db)^2 + (dc)^2; \quad (ds)^2 = (dr)^2 + r^2(d\theta)^2 + r^2 \sin^2\theta(d\phi)^2.$$

On replacing $d\,{}^k a$ by its equivalent: $d({}^k a) = {}^k a_{,\sigma}\, dx^\sigma$ we may express $(ds_0)^2$ in terms of the differentials of the final coördinates x^r:

$$(ds_0)^2 = {}_{\alpha\beta}c(d\,{}^\alpha a)(d\,{}^\beta a) = h_{\sigma\tau}\, dx^\sigma\, dx^\tau$$

where

$$h_{pq} = {}_{\alpha\beta}c({}^\alpha a_{,p})({}^\beta a_{,q}).$$

In a similar manner we may express $(ds)^2$ in terms of the differentials of the initial coördinates ${}^r a$:

$$(ds)^2 = g_{\alpha\beta}\, dx^\alpha\, dx^\beta = {}_{\alpha\beta}k(d\,{}^\alpha a)(d\,{}^\beta a)$$

where

$${}_{pq}k = g_{\alpha\beta}({}_{p,}x^\alpha)({}_{q,}x^\beta).$$

We obtain thus two equivalent expressions for the difference $(ds)^2 - (ds_0)^2$ of the initial and final squared elements of arc length:

$$(ds)^2 - (ds_0)^2 = 2\epsilon_{\alpha\beta}\, dx^\alpha\, dx^\beta = 2\, {}_{\alpha\beta}\eta(d\,{}^\alpha a)(d\,{}^\beta a)$$

where

$$2\epsilon_{pq} = g_{pq} - h_{pq}; \qquad 2\, {}_{pq}\eta = {}_{pq}k - {}_{pq}c$$

For a displacement in which lengths are preserved the difference of the squared elements of arc length is zero (identically in the differentials $d\,{}^r a$, or, equivalently, the differentials dx^r) and so the quantities ϵ_{pq} and ${}_{pq}\eta$ are zero for such a *rigid* displacement. In general we regard the quantities ϵ_{pq} or ${}_{pq}\eta$ as descriptive of the strain or deformation and we thus have two methods of describing the strain:

(a) the description, by means of the quantities ϵ_{pq}, in which the final coördinates x^r are adopted as the independent variables in terms of which the description is made and

(b) the description, by means of the quantities ${}_{pq}\eta$, in which the initial coördinates ${}^r a$ are adopted as the independent variables in terms of which the description is made.

In the terminology of the introduction these are, respectively, the Eulerian and Lagrangian descriptions of the strain. We shall also refer to them as the tensor and scalar descriptions, respectively; and shall call the quantities ϵ_{pq} the tensor strain-components and the quantities ${}_{pq}\eta$ the scalar strain-components. The quantities ϵ_{pq} are the covariant components of the strain tensor; this tensor may also be presented in mixed form, or in contravariant form, by means of the formulae

$$\epsilon_s{}^r = g^{ra}\,\epsilon_{as}\,; \qquad \epsilon^{rs} = g^{ra}\,\epsilon_a{}^s$$

where g^{rs} is the reciprocal of g_{rs} : $g^{ra}\,g_{as} = \delta_s{}^r$. Similarly it is convenient to introduce other (but equivalent) scalar descriptions of the strain by means of the formulae:

$${}^p{}_q\eta = ({}^{pa}c)\,({}_{aq}\eta)\,; \qquad {}^{pq}\eta = ({}^{pa}c)\,({}^q{}_a\eta)$$

where ${}^{pq}c$ is the reciprocal of ${}_{pq}c$: $({}^{pa}c)\,({}_{aq}c) = {}^p{}_q\delta$. In the technical language of tensor analysis we may say that we use the tensor g_{pq}, and its reciprocal g^{pq}, for stepping labels down and up upon tensor quantities; and the matrix ${}_{pq}c$, and its reciprocal ${}^{pq}c$, for stepping labels down and up upon scalar sets.

When rectangular Cartesian coördinates, relative to the same reference frame, are used for both the initial and final positions the tensor strain components take the form

$$2\epsilon_{xx} = 1 - \left(\frac{\partial a}{\partial x}\right)^2 - \left(\frac{\partial b}{\partial x}\right)^2 - \left(\frac{\partial c}{\partial x}\right)^2;$$
$$2\epsilon_{yz} = -\frac{\partial a}{\partial y}\frac{\partial a}{\partial z} - \frac{\partial b}{\partial y}\frac{\partial b}{\partial z} - \frac{\partial c}{\partial y}\frac{\partial c}{\partial z} \quad \text{etc.,}$$

whilst the scalar strain components are given by

$$2{}_{aa}\eta = \left(\frac{\partial x}{\partial a}\right)^2 + \left(\frac{\partial y}{\partial a}\right)^2 + \left(\frac{\partial z}{\partial a}\right)^2 - 1\,;$$
$$2{}_{bc}\eta = \frac{\partial x}{\partial b}\frac{\partial x}{\partial c} + \frac{\partial y}{\partial b}\frac{\partial y}{\partial c} + \frac{\partial z}{\partial b}\frac{\partial z}{\partial c} \quad \text{etc.}$$

On denoting the displacement vector $(x-a, y-b, z-c)$ by (u, v, w) we have

$$\epsilon_{xx} = \frac{\partial u}{\partial x} - \tfrac{1}{2}\left\{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial x}\right)^2\right\};$$
$$\epsilon_{yz} = \tfrac{1}{2}\left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right) - \tfrac{1}{2}\left\{\frac{\partial u}{\partial y}\frac{\partial u}{\partial z} + \frac{\partial v}{\partial y}\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\frac{\partial w}{\partial z}\right\} \quad \text{etc.,}$$
$${}_{aa}\eta = \frac{\partial u}{\partial a} + \tfrac{1}{2}\left\{\left(\frac{\partial u}{\partial a}\right)^2 + \left(\frac{\partial v}{\partial a}\right)^2 + \left(\frac{\partial w}{\partial a}\right)^2\right\};$$
$${}_{bc}\eta = \tfrac{1}{2}\left(\frac{\partial v}{\partial c} + \frac{\partial w}{\partial b}\right) + \tfrac{1}{2}\left\{\frac{\partial u}{\partial b}\frac{\partial u}{\partial c} + \frac{\partial v}{\partial b}\frac{\partial v}{\partial c} + \frac{\partial w}{\partial b}\frac{\partial w}{\partial c}\right\} \quad \text{etc.}$$

In the classical infinitesimal theory the partial derivatives $\partial u/\partial x, \cdots, \partial u/\partial a, \cdots$ are regarded as infinitesimal and so we may put

$$\epsilon_{xx} = \frac{\partial u}{\partial x}; \qquad \epsilon_{yz} = \tfrac{1}{2}\left(\frac{\partial v}{\partial z} + \frac{\partial w}{\partial y}\right) \text{ etc.,}$$

$$_{aa}\eta = \frac{\partial u}{\partial a}; \qquad _{bc}\eta = \tfrac{1}{2}\left(\frac{\partial v}{\partial c} + \frac{\partial w}{\partial b}\right) \text{ etc.}$$

Since

$$\frac{\partial u}{\partial a} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial a} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial a} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial a} = \frac{\partial u}{\partial x}\left(1 + \frac{\partial u}{\partial a}\right) + \frac{\partial u}{\partial y}\frac{\partial v}{\partial a} + \frac{\partial u}{\partial z}\frac{\partial w}{\partial a}$$

we have, to the degree of approximation contemplated by the infinitesimal theory, $\partial u/\partial a = \partial u/\partial x$ etc., and there is no distinction between the tensor strain components and the scalar strain components.

Relation between the elements of volume. A region occupied by the medium in its initial position is described by setting the coördinates $^r a$ functions of some convenient three independent variables and then the initial element of volume dV_0 is given by the formula $dV_0 = \sqrt{c}\,|\,d(a)\,|$ where c denotes the determinant of the matrix $_{pq}c$ and $|\,d(a)\,|$ denotes the numerical value of the product of the differentials of the independent variables by the Jacobian determinant of the three coördinates $^r a$ relative to the independent variables. Similarly dV, the element of volume occupied by the *same* particles when in the strained position, is given by the formula $dV = \sqrt{g}\,|\,d(x)\,|$ where g denotes the determinant of the tensor g_{rs}. Now the relation

$$_{\alpha\beta}c\,(d^{\alpha}a)\,(d^{\beta}a) = (ds_0)^2 = h_{\alpha\beta}\, dx^{\alpha}\, dx^{\beta}$$

implies

$$\sqrt{c}\,|\,d(a)\,| = \sqrt{h}\,|\,d(x)\,|$$

where h denotes the determinant of the tensor h_{rs}. On writing this tensor in its mixed form: $h_{rs} = g_{r\alpha}\, h_s{}^{\alpha}$ and using the theorem that the determinant of the product of two matrices equals the product of the determinants of the two factors we find $h = g\,\det(h_s{}^r)$ and so

$$\begin{aligned} dV_0/dV &= \sqrt{c}\,|\,d(a)\,|/\sqrt{g}\,|\,d(x)\,| \\ &= \sqrt{h}\,|\,d(x)\,|/\sqrt{g}\,|\,d(x)\,| = \sqrt{h/g} = \sqrt{\det(h_s{}^r)}. \end{aligned}$$

On writing the relation $h_{pq} = g_{pq} - 2\epsilon_{pq}$ in its mixed form $h_p{}^q = \delta_q{}^p - 2\epsilon_q{}^p$ we find $\det(h_s{}^r) = 1 - 2I_1 + 4I_2 - 8I_3$ where

$$I_1 = \epsilon_{\alpha}{}^{\alpha}; \qquad I_2 = \frac{1}{2\,!}\,\delta^{\alpha_1\alpha_2}_{\beta_1\beta_2}\epsilon_{\alpha_1}{}^{\beta_1}\epsilon_{\alpha_2}{}^{\beta_2}; \qquad I_3 = \frac{1}{3\,!}\,\delta^{\alpha_1\alpha_2\alpha_3}_{\beta_1\beta_2\beta_3}\epsilon_{\alpha_1}{}^{\beta_1}\epsilon_{\alpha_2}{}^{\beta_2}\epsilon_{\alpha_3}{}^{\beta_3}$$

$= \epsilon^2_2\epsilon^3_3 + \epsilon^3_3\epsilon^1_1$ $\qquad = |\epsilon^i_j|$

$+ \epsilon^1_1\epsilon^2_2 - \epsilon^2_3\epsilon^3_2$

$- \epsilon^3_1\epsilon^1_3 - \epsilon^1_2\epsilon^2_1$

are the three strain invariants. They are, respectively, the sum of the diagonal elements, the sum of the principal two rowed minors, and the determinant of the matrix $\epsilon_q{}^p$ which presents the strain tensor in mixed form. When the strain is homogeneous i. e. does not vary from point to point of the medium we may replace the ratio dV_0/dV by the ratio V_0/V where V_0 is the initial volume of any portion of the medium and V is the volume occupied in the strained position, by the particles which initially occupied the volume V_0. We have, then, for a homogeneous strain the relation

$$V_0/V = \sqrt{\det(\delta_s{}^r - 2\epsilon_s{}^r)} = \sqrt{1 - 2I_1 + 4I_2 - 8I_3}.$$

For the special case of a homogeneous strain which is also at each point isotropic (which will be the case in an isotropic medium subjected to uniform hydrostatic pressure) the strain tensor will be a scalar tensor: $\epsilon_s{}^r = \epsilon\delta_s{}^r$ and we have the relatively simple relation

$$V_0/V = (1 - 2\epsilon)^{3/2}.$$

For the classical infinitesimal theory this reduces, as is at once seen on writing $(1 - 2\epsilon)^{3/2} = 1 - 3\epsilon + \cdots$, to $V_0/V = 1 - 3\epsilon$ or

$$\epsilon = \tfrac{1}{3} \cdot \frac{V - V_0}{V} = \tfrac{1}{3}\frac{\Delta V}{V} = \tfrac{1}{3}\frac{\Delta V}{V_0}$$

(to the degree of approximation contemplated by the theory). The exact relation, valid for a finite strain, to which the relation just written is an approximation is

$$\epsilon = \tfrac{1}{2}\{1 - (V_0/V)^{2/3}\}.$$

The mathematical description of a homogeneous strain is that the strain tensor should be constant *in the tensor sense* i. e. its absolute or covariant derivative vanishes. This implies that the absolute derivatives of the invariants I_1, I_2, I_3 (or what is the same thing, since these are scalars, their space derivatives) vanish so that dV_0/dV is a numerical constant. The hypothesis $\epsilon_{pq,r} = 0$ (where as is usual labels following a comma indicate covariant differentiation) implies $h_{pq,r} = 0$ and since $h_{pq} = ({}_{\alpha\beta}c)\, {}^{\alpha}a_{,p}\, {}^{\beta}a_{,q}$ an easy calculation yields ${}^{k}a_{,pq} + ({}^{k}{}_{\alpha\beta}\Gamma)\, {}^{\alpha}a_{,p}\, {}^{\beta}a_{,q} = 0$. In particular when the a's and x's are rectangular Cartesian coördinates so that ${}^{k}a_{,pq} = \partial^2\, {}^{k}a/\partial x^p\, \partial x^q$ and ${}^{k}{}_{rs}\Gamma = 0$ the ${}^{k}a$ must be linear functions of the x^r. Conversely if this is the case the strain components are numerical constants and the strain is homogeneous.

The variation of the strain tensor. In order to introduce the concept of virtual work and thereby express the conditions for equilibrium of the strained medium we must adopt the dynamic as opposed to the static viewpoint. In other words instead of regarding the strained position of the medium as something fixed and final we must regard it as capable of variation. To do this

2

conveniently, from the mathematical viewpoint, we conceive of the final coördinates x^s as depending not only upon the initial coördinates ${}^r a$ but also upon an accessory parameter θ (which could, in hydrodynamics, conveniently be taken as the time variable). We shall denote differentials with respect to the parameter θ by the symbol D:

$$Dx^s = \frac{\partial x^s}{\partial \theta} d\theta$$

it being understood that in the partial differentiation with respect to θ the coördinates ${}^r a$ are kept constant (i. e. the D denotes the substantial or particle differentiation of hydrodynamics). If we have any tensor function $f^{\cdots}_{\cdots\cdot}$ of the coördinates x^r we shall denote by $\delta f^{\cdots}_{\cdots\cdot}$ the tensor of the same type defined by the rule

$$\delta f^{\cdots}_{\cdots\cdot} = f^{\cdots}_{\cdots\cdot, a} Dx^a.$$

When the coördinates x^r are Cartesian covariant differentiation is merely ordinary space differentiation of the tensor components and so $\delta f^{\cdots}_{\cdots\cdot} = Df^{\cdots}_{\cdots\cdot}$. Even when the coördinates are not Cartesian this relation holds for any scalar function: $\delta f = Df$ since $f_{,r} = \partial f / \partial x^r$. We effect a small economy of notation by defining δx^r by the relation $\delta x^r = Dx^r$ and we refer to the contravariant vector δx^r as the virtual displacement vector. Before proceeding to the necessary calculation of the variation of the strain tensor and of its scalar components it must be clearly understood that the variation of any scalar function of the ${}^r a$ is zero; e. g. $\delta_{pq} c = 0$. This result is a trivial consequence of the definition of the δ symbol for in the differentiation with respect to θ the coördinates ${}^r a$ are held constant.

Owing to the independence of the variables ${}^r a$ and θ differentiations with respect to them are interchangeable as to order (it being supposed that the second order derivatives involved exist and are continuous). Hence

$$D({}_{s,}x^k) = \frac{\partial}{\partial {}^s a} \delta x^k = \frac{\partial}{\partial x^a} (\delta x^k) \cdot ({}_{s,}x^a)$$

and on multiplication by $d^s a$ (which is independent of θ) and summation with respect to s we find

$$D(dx^k) = \frac{\partial}{\partial x^a} (\delta x^k) \cdot dx^a.$$

This implies the tensor equation

$$\delta(dx^k) = (\delta x^k)_{,a}\, dx^a$$

since it is the form to which the tensor equation reduces when the coördinates

x^k are Cartesian. Since the variation of the metrical tensor is zero: $\delta g_{rs} = 0$, the tensor equation just written may be put in the equivalent form

$$\delta(dx_k) = (\delta x_k)_{,a}\, dx^a.$$

Since $d({}^r a) = {}^r a_{,a}\, dx^a$ and $\delta d({}^r a) = 0$ we have

$$\delta({}^r a_{,a})\, dx^a = -\, {}^r a_{,a}\, \delta(dx^a) = -\, {}^r a_{,a} (\delta x^a)_{,\beta}\, dx^\beta$$

and since this must hold for arbitrary dx^r we must have

$$\delta({}^r a_{,k}) = -\, {}^r a_{,a} (\delta x^a)_{,k}.$$

From this expression and the relation $h_{pq} = {}_{a\beta}c({}^a a_{,p})({}^\beta a_{,q})$ we can read off at once the formula for the variation of h_{pq} (or, what is the same thing, of $-2\epsilon_{pq}$). We find

$$\begin{aligned} \delta h_{pq} &= {}_{a\beta}c(\delta^a a_{,p} \cdot {}^\beta a_q + {}^a a_{,p}\, \delta^\beta a_{,q}) \\ &= -\, {}_{a\beta}c\{{}^a a_{,\tau} (\delta x^\tau)_{,p}\, {}^\beta a_q + ({}^a a_{,p}) {}^\beta a_{,\tau} (\delta x^\tau)_{,q}\} \\ &= -\, h_{\tau q} (\delta x^\tau)_{,p} - h_{p\tau} (\delta x^\tau)_{,q} \\ &= -\, h_q{}^\tau (\delta x_\tau)_{,p} - h_p{}^\tau (\delta x_\tau)_{,q}. \end{aligned}$$

This will be seen to be the significant formula for our later purpose. On writing $h_{pq} = g_{pq} - 2\epsilon_{pq}$ it may be written in the equivalent form

$$\delta\epsilon_{pq} = \tfrac{1}{2}\{(\delta x_q)_{,p} + (\delta x_p)_{,q}\} - \{\epsilon_q{}^\tau (\delta x_\tau)_{,p} + \epsilon_p{}^\tau (\delta x_\tau)_{,q}\}.$$

For the classical infinitesimal theory it is allowable to write

$$\delta\epsilon_{pq} = \tfrac{1}{2}\{(\delta x_q)_{,p} + (\delta x_p)_{,q}\}$$

and it is the difference between this and the exact expression just furnished that makes it incorrect, as stated in the introduction, to write the stress tensor as the gradient with respect to the strain tensor of the elastic energy density.

Criterion for a rigid virtual displacement. A virtual displacement is said to be rigid when $\delta(ds^2) = 0$. Since $\delta(ds_0)^2 = 0$ and

$$(ds)^2 - (ds_0)^2 = 2\epsilon_{a\beta}\, dx^a\, dx^\beta$$

an equivalent description of a rigid virtual displacement is

$$\delta(\epsilon_{a\beta}\, dx^a\, dx^\beta) = 0.$$

One must be on one's guard against the error of supposing that (because in a rigid displacement $\epsilon_{pq} = 0$) in a virtual rigid displacement the strain tensor

ϵ_{pq} is constant: $\delta\epsilon_{pq} = 0$. The evident fallacy in such a guess being the neglect of terms involving $\delta(dx^r)$ which quantities are not in general zero in a rigid displacement. On the other hand the scalar components $_{pq}\eta$ of the strain tensor are constant in a rigid virtual displacement; for the criterion for a rigid virtual displacement may be put in the form $\delta(_{\alpha\beta}\eta \, d\,^\alpha a \, d\,^\beta a) = 0$ and this is equivalent to $\delta \, _{\alpha\beta}\eta (d\,^\alpha a)(d\,^\beta a) = 0$ since $d\,^r a$ is independent of θ. Since this equation must hold for arbitrary $d\,^r a$ we must have $\delta_{pq}\eta = 0$ and this necessary condition is clearly sufficient. For any virtual displacement, rigid or not, we have

$$\begin{aligned} \delta(ds)^2 &= \delta(g_{\alpha\beta}\, dx^\alpha\, dx^\beta) = \delta(dx^\alpha\, dx_\alpha) \\ &= dx^\alpha (\delta x_\alpha)_{,\beta}\, dx^\beta + (\delta x^\alpha)_{,\beta}\, dx^\beta\, dx_\alpha \\ &= \{(\delta x_\alpha)_{,\beta} + (\delta x_\beta)_{,\alpha}\} dx^\alpha\, dx^\beta \end{aligned}$$

Interchange dummies

so that the criterion for a rigid virtual displacement may be put in the form $(\delta x_p)_{,q} + (\delta x_q)_{,p} = 0$ (equations of Killing). Amongst the possible rigid virtual displacements are those for which $(\delta x_p)_{,q} = 0$; these are the virtual translations which appear, when the coördinates x_p are Cartesian, in the form $\delta x_p =$ constant. Since $\delta(ds)^2 = 2\delta(_{\alpha\beta}\eta \, d\,^\alpha a \, d\,^\beta a)$ we have, for any virtual displacement whatsoever,

$$2\delta \, _{\alpha\beta}\eta \cdot d\,^\alpha a \cdot d\,^\beta a = \{(\delta x_\alpha)_{,\beta} + (\delta x_\beta)_{,\alpha}\} dx^\alpha\, dx^\beta$$

implying

$$\delta \, _{pq}\eta = \tfrac{1}{2}\{(\delta x_\alpha)_{,\beta} + (\delta x_\beta)_{,\alpha}\}(_{p,}x^\alpha)(_{q,}x^\beta).$$

2. The stress-tensor and the virtual work of the applied forces acting on the medium. Let us consider any portion of our elastic medium which is bounded by a closed surface and denote by S_0 and S the initial and final positions, respectively, of this bounding surface. The coördinates $^r a$ of the initial position of any particle of this bounding surface (and equally the coördinates x^s of the final position of such a particle) are functions of two independent parameters and the surface element of S may be described by means of the covariant vector $dS_r = \sqrt{g}\, d(x^p, x^q)$ (where p, q, r is an even permutation of the natural order 1, 2, 3 and $d(x^p, x^q)$ denotes the product of the differentials of the two independent variables times the Jacobian of the two coördinates x^p and x^q relative to these independent variables). When the coördinates x^r are rectangular Cartesian

$$dS_x = d(y, z); \qquad dS_y = d(z, x); \qquad dS_z = d(x, y).$$

Denoting by dS the magnitude of this typical surface element—so that

$$(dS)^2 = g^{\alpha\beta}\, dS_\alpha\, dS_\beta$$

—the stress tensor T^{rs} is a linear vector function which associates with each surface element dS_r a stress vector F^r by means of the formula $F^r dS = T^{ra} dS_a$. The virtual work of the stresses across the boundary S is accordingly

$$\int_S (F^\beta \, dS) \delta x_\beta = \int_S T^{\beta a} \, dS_a \, \delta x_\beta$$

and this is equivalent to the volume integral

$$\int_V (T^{\beta a} \, \delta x_\beta)_{,a} \, dV$$

extended over the volume V bounded by S. If there are mass forces—M^r per unit mass—acting on the medium the virtual work of these is $\int_V \rho M^\beta \, \delta x_\beta \, dV$, where ρ is the mass density, and so the virtual work of all the forces acting on any portion of the medium is

$$\int_V \{ (T^{\beta a}{}_{,a} + \rho M^\beta) \, \delta x_\beta + T^{\beta a} (\delta x_\beta)_{,a} \} dV.$$

We now make the physical assumption (criterion of equilibrium) that this virtual work is zero for any *rigid* virtual displacement. Amongst these rigid virtual displacements are the translations which are characterized by the tensor equation $(\delta x_p)_{,q} = 0$ and so we must have

$$\int_V (T^{\beta a}{}_{,a} + \rho M^\beta) \delta x_\beta \, dV = 0;$$

since δx_r may be assigned arbitrarily at a given point and since the volume V of integration is arbitrary this forces

$$T^{ra}{}_{,a} + \rho M^r = 0.$$

Consequently the virtual work of all the forces (mass and surface) acting upon any portion of the medium in any virtual displacement whatever is given by the expression

$$\text{Virtual work} = \int_V T^{\beta a} (\delta x_\beta)_{,a} \, dV.$$

Furthermore since this must vanish for any rigid virtual displacement i. e. for any virtual displacement for which $(\delta x_p)_{,q} + (\delta x_q)_{,p} = 0$ the stress tensor must be symmetric: $T^{sr} = T^{rs}$. This relation enables us to write the expression for the virtual work of all the forces acting upon the medium in any virtual displacement whatever in the form:

$$\text{Virtual work} = \tfrac{1}{2}\int_V T^{\alpha\beta}\{(\delta x_\alpha)_{,\beta} + (\delta x_\beta)_{,\alpha}\}dV.$$

In the classical infinitesimal theory this may be written as $\int T^{\alpha\beta}\,\delta\epsilon_{\alpha\beta}\,dV$ but such an approximation is not legitimate in the finite theory.

3. The elastic potential and its connection with the stress tensor. We now consider an element of volume dV of the medium in its strained position and denote by ρ the density of the matter occupying the element of volume dV so that the element of mass is $dm = \rho\,dV$. The principle of conservation of mass is expressed by the formula:

$$\delta(dm) = \delta(\rho\,dV) = 0.$$

In order to apply the fundamental energy-conservation law of thermodynamics we denote by T the temperature of the element of mass dm; by σ the entropy density (per unit mass) so that the entropy of the mass dm is $\sigma\,dm = \rho\sigma\,dV$; and by udm the internal energy of the mass dm. Then the principle of thermodynamics to which we have referred says that $T\delta(\sigma\,dm) = \delta(udm)$—Virtual work of all forces acting on dm. On introducing the free-energy density $\phi = u - T\sigma$ and availing ourselves of the principle of conservation of mass: $\delta\,dm = 0$ we find, on integrating over any portion V of the strained medium,

$$\int_V \delta\phi\,dm = \int_V T^{\alpha\beta}(\delta x_\alpha)_{,\beta}\,dV - \int_V \sigma\,dm\cdot\delta T.$$

On writing $dm = \rho\,dV$ and observing that this relation must hold for arbitrary volumes V we obtain the fundamental formula connecting the elastic potential ϕ with the stress tensor T^{rs}:

$$\rho\,\delta\phi = T^{\alpha\beta}(\delta x_\alpha)_{,\beta} - \rho\sigma\,\delta T.$$

In order to derive from this a connection between the stress tensor T^{rs} and the strain tensor ϵ_{rs} we must make some hypothesis concerning the function ϕ. We shall assume that it depends only on the three gradient vectors ${}^ra_{,s}$ (or, equivalently, on the reciprocal set ${}_{r,}x^s$) it being understood that either set may appear both covariantly and contravariantly. In other words we shall assume that ϕ is a function of the vectors ${}^ra_{,s}$, the metrical tensor g_{rs}, the scalar quantities ${}_{pq}c$ and the temperature T. We shall confine our attention in what follows to isothermal variations so that T is a constant parameter in ϕ to which attention need not be explicitly directed. Then $\delta\phi$ must be zero in any (isothermal) rigid virtual displacement; in other words $\dfrac{\partial\phi}{\partial({}^\alpha a_{,\beta})}\,\delta({}^\alpha a_{,\beta}) = 0.$

provided $(\delta x_p)_{,q} + (\delta x_q)_{,p} = 0$. On inserting the value $-({}^r a_{,a})(\delta x^a)_{,k}$ given in .1 for $\delta({}^r a_{,k})$ we see that the tensor $\frac{\partial \phi}{\partial({}^a a_{,p})}({}^a a^{,q})$ must be symmetric in p and q $({}^r a^{,q} = g^{qa}({}^r a_{,a}))$. Hence the function ϕ of the nine variables ${}^r a_{,s}$ is conditioned by the three linear, homogeneous, first order, partial differential equations:

$$\frac{\partial \phi}{\partial({}^a a_{,p})}\, {}^a a^{,q} = \frac{\partial \phi}{\partial({}^a a_{,q})}\, {}^a a^{,p}.$$

These equations are readily seen to form a complete system (the commutator of any two being the third) and so the general solution of them is a function of $6 = 9 - 3$ independent solutions. Since $\delta\, {}_{pq}\eta = 0$ in any rigid virtual displacement we know that the six functions ${}_{pq}\eta$ satisfy the three partial differential equations just written. Hence ϕ can involve the vectors ${}^r a_{,s}$ only through the scalar functions ${}_{pq}\eta$: $\phi = \phi({}_{pq}\eta, {}_{pq}c, T)$. The quantities g_{rs} cannot occur in the expression for ϕ since ϕ is a scalar and the only scalar functions of the metrical tensor g_{rs} are numerical (e. g. $g_a{}^a = \delta_a{}^a = 3$). Now under a transformation of coördinates ${}^r a$ the quantities ${}_{pq}\eta, {}_{pq}c$ (which are scalars under transformation of coördinates x^s) transform as covariant tensors. We say that the medium is *isotropic* when the elastic-energy density ϕ is unaffected by a transformation of the coördinates ${}^r a$; for instance when the coördinates ${}^r a$ are rectangular Cartesian an arbitrary rotation of the reference frame must leave the function ϕ unaffected. In order that this may be the case ϕ must involve ${}_{pq}\eta, {}_{pq}c$ only through the "invariants" (under transformation of the coördinates ${}^r a$):

$$J_1 = {}^a{}_a\eta\,; \quad J_2 = \tfrac{1}{2}\,\delta^{a_1 a_2}_{\beta_1\beta_2}({}^{\beta_1}{}_{a_1}\eta)({}^{\beta_2}{}_{a_2}\eta)\,; \quad J_3 = \frac{1}{3\,!}\,\delta^{a_1 a_2 a_3}_{\beta_1\beta_2\beta_3}({}^{\beta_1}{}_{a_1}\eta)({}^{\beta_2}{}_{a_2}\eta)({}^{\beta_3}{}_{a_3}\eta).$$

These quantities J_1, J_2, J_3 are, respectively, the sum of the diagonal elements, the sum of the principal two-rowed minors, and the determinant of the matrix ${}^p{}_q\eta$. Hence, for an isotropic medium: $\phi = \phi(J_1, J_2, J_3, T)$. We prove in the appendix that the quantities J_1, J_2, J_3 are functions of the three invariants I_1, I_2, I_3 of the stress tensor $\epsilon_s{}^r$ and so we may write, for an isotropic medium,

$$\phi = \phi(I_1, I_2, I_3, T)$$

i. e. ϕ is a function of the components of the strain tensor $\epsilon_s{}^r$ and T. Conversely if we make the hypothesis that ϕ is a function of T and of the strain components (tensor) $\epsilon_s{}^r$ *alone* this implies that ϕ is isotropic; for the only scalar functions of $\epsilon_s{}^r$ are functions of its invariants I_1, I_2, I_3. In order to prevent possible misunderstanding of this remark [in view of the fact that in

the classical treatment of elasticity ϕ is taken for crystalline (= non-isotropic) media as a quadratic function of the strain components] we may say that it is tacitly understood in this classical procedure that a special privileged reference frame, determined by the axes of the crystal, has been chosen. The coefficients of the quadratic form are accordingly not scalar quantities but constitute a tensor which depends on the orientation of the crystalline axes.

The fundamental stress-strain relations for an isotropic medium. As we have just seen the elastic energy density ϕ is, for an isotropic medium, a function of the tensor strain components $\epsilon_s{}^r$, involving these through the strain invariants I_1, I_2, I_3. It will be, for the moment, more convenient to regard ϕ as a function of the covariant strain components ϵ_{rs} ($\epsilon_s{}^r = g^{ra}\epsilon_{as}$!) and we shall, if necessary, symmetrize its formal expression; i. e. we shall replace each ϵ_{rs}, wherever it occurs in the expression for ϕ, by its equivalent: $\epsilon_{rs} = \frac{1}{2}(\epsilon_{rs} + \epsilon_{sr})$. Denoting by $\partial\phi/\partial\epsilon_{rs}$ the partial derivative of ϕ with respect to ϵ_{rs} all the other ϵ_{rs} (including ϵ_{sr}) being held constant in the differentiation (so that in this formal differentiation no attention is paid to the symmetry relations $\epsilon_{sr} = \epsilon_{rs}$) it follows that $\partial\phi/\partial\epsilon_{rs} = \partial\phi/\partial\epsilon_{sr}$. We have seen in 1 that

$$\delta\epsilon_{pq} = -\tfrac{1}{2}\delta h_{pq} = \tfrac{1}{2}\{h_q{}^\tau(\delta x_\tau)_{,p} + h_p{}^\tau(\delta x_\tau)_{,q}\}$$

where $h_{pq} = g_{pq} - 2\epsilon_{pq}$, and hence

$$\delta\phi = \frac{\partial\phi}{\partial\epsilon_{\alpha\beta}}\,\delta\epsilon_{\alpha\beta} = \frac{1}{2}\,\frac{\partial\phi}{\partial\epsilon_{\alpha\beta}}\,\{h_\beta{}^\tau(\delta x_\tau)_{,\alpha} + h_\alpha{}^\tau(\delta x_\tau)_{,\beta}\} = \frac{\partial\phi}{\partial\epsilon_{\alpha\beta}}\,h_\beta{}^\tau(\delta x_\tau)_{,\alpha}$$

since $\partial\phi/\partial\epsilon_{rs} = \partial\phi/\partial\epsilon_{sr}$. Since in an isothermal virtual displacement $\rho\,\delta\phi = T^{\alpha\beta}(\delta x_\alpha)_{,\beta}$ it follows that, for an isotropic medium,

$$\rho\,\frac{\partial\phi}{\partial\epsilon_{\alpha\beta}}\,h_\beta{}^\tau(\delta x_\tau)_{,\alpha} = T^{\alpha\beta}(\delta x_\alpha)_{,\beta}.$$

Since the virtual displacement is arbitrary the components $(\delta x_p)_{,q}$ of the space derivative of the virtual displacement vector may be assigned arbitrary values at any point x^r and so we must have

$$T^{rs} = \rho\,\frac{\partial\phi}{\partial\epsilon_{s\beta}}\,h_\beta{}^r = \rho\left(\frac{\partial\phi}{\partial\epsilon_{sr}} - 2\epsilon_\beta{}^r\,\frac{\partial\phi}{\partial\epsilon_{s\beta}}\right).$$

That these equations imply $T^{sr} = T^{rs}$ follows from the fact that ϕ involves the strain components ϵ_{rs} only through the strain invariants I_1, I_2, I_3. For example $I_1 = \epsilon_\alpha{}^\alpha = g^{\alpha\beta}\epsilon_{\alpha\beta}$ so that

$$\epsilon_{\beta}{}^{r} \frac{\partial I_1}{\partial \epsilon_{s\beta}} = \epsilon_{\beta}{}^{r} g^{s\beta} = \epsilon^{rs} = \epsilon_{\beta}{}^{s} \frac{\partial I_1}{\partial \epsilon_{r\beta}}.$$

The formulae just obtained are fundamental but it is frequently more convenient to present the stress tensor in its mixed form $T_s{}^r$. Since

$$\epsilon_s{}^r = g^{ra}\,\epsilon_{as}\,; \quad \frac{\partial \phi}{\partial \epsilon_{rs}} = g^{ra}\,\frac{\partial \phi}{\partial \epsilon_s{}^a}$$

and so our formulae appear as

$$T_s{}^r = \rho\left(\frac{\partial \phi}{\partial \epsilon_r{}^s} - 2\epsilon_\beta{}^r\,\frac{\partial \phi}{\partial \epsilon_\beta{}^s}\right).$$

The result $dV_0/dV = \sqrt{1 - 2I_1 + 4I_2 - 8I_3}$, obtained in **1**, may be written in the equivalent form $\rho = \rho_0\sqrt{1 - 2I_1 + 4I_2 - 8I_3}$ (since the principle of conservation of mass implies $\rho\, dV = \rho_0\, dV_0$). In the classical (infinitesimal) theory the strain invariants I_1, I_2, I_3 are infinitesimal quantities of the first, second and third orders of magnitude respectively. We may therefore write, to a first approximation, $\rho = \rho_0$ and to a second approximation $\rho = \rho_0(1 - I_1)$. Keeping only the first approximation our fundamental stress strain relations reduce to

$$T_s{}^r = \rho_0\,\frac{\partial \phi}{\partial \epsilon_r{}^s} = \frac{\partial \phi'}{\partial \epsilon_r{}^s}; \qquad \phi' = \rho_0 \phi.$$

These are the basic formulae (expressing Hooke's law) of the classicical theory, ϕ' being the elastic energy per unit initial volume (or, what is the same thing to the degree of approximation contemplated by the classical theory, per unit final volume).

Even in the case of finite strains the strain invariants are relatively small. The two cases to which we shall devote some attention in detail in the present paper are:

(a) uniform hydrostatic pressure; here the stress and strain tensors are scalar tensors: $T_s{}^r = -p\delta_s{}^r$; $\epsilon_s{}^r = -f\delta_s{}^r$ (where p is what is commonly called pressure) and $I_1 = -3f$, $I_2 = 3f^2$, $I_3 = -f^3$. We shall discuss a little later the agreement of the finite theory with the results of experiments by Bridgman (**10**) on the compressibility of sodium under pressures ranging from 2,000 atmospheres to 20,000 atmospheres. In these experiments the ratio $(V_0 - V)/V_0$ varied from .030 at a pressure of 2,000 atm., to .189 at a pressure of 20,000 atm. Since $V_0/V = \sqrt{\det(\delta_s{}^r - 2\epsilon_s{}^r)} = (1 + 2f)^{3/2}$ the corresponding values of f are .010 and .075 respectively; hence I_1 varied between $-.030$ and $-.225$, I_2 varied between .001 and .0169 whilst I_3 varied between $-.000001$ and $-.00042$.

(b) uniform linear stress (Young's modulus experiment). Here the stress tensor is scalar with two components zero whilst the strain tensor is scalar with two components equal:

$$T_x^x = T_y^y = 0; \qquad \epsilon_x^x = \epsilon_y^y = -\sigma\epsilon_z^z$$

where σ, Poisson's ratio, has a value $< .5$. From the formula giving the volume change we have $V_0/V = (1 + 2\sigma\epsilon_z^z)\sqrt{1 - 2\epsilon_z^z}$ so that ϵ_z^z cannot surpass the value .5. Hence $I_1 = (1 - 2\sigma)\epsilon_z^z$ is a small fraction even for very large strains and $I_2 = (\sigma^2 - 2\sigma)(\epsilon_z^z)^2$, $I_3 = \sigma^2(\epsilon_z^z)^3$ are smaller.

We may, therefore, even for large strains hope to secure good agreement with experiment by expanding $\phi(\epsilon_s^r)$ as a power series in the strain components and neglecting terms of orders of magnitude greater than an agreed upon order (say the second, third, etc.). In the classical, infinitesimal, theory the agreed upon order is the second. We shall agree upon the third but we call explicit attention at this point to the fact that our theory gives remarkable agreement with experimental results even if we do not introduce any more constants (i. e. coefficients in the expansion of the elastic energy density ϕ) than those (two in number) introduced in the infinitesimal theory. In fact in the case of the compressibility experiments the two constants combine into a single one so that a *one* constant formula suffices to predict to a high degree of accuracy the connection between pressure and volume over the extensive range from 2,000 to 20,000 atmospheres. It is clear that an additive constant in ϕ is of no significance since ϕ enters our fundamental equations only through its partial derivatives. In order to keep as closely as possible to the notations of the classical (infinitesimal) theory we shall expand $\rho_0\phi$ instead of ϕ:

$$\rho_0\phi = \alpha I_1 + \frac{\lambda + 2\mu}{2} I_1^2 - 2\mu I_2 + lI_1^3 + mI_1I_2 + nI_3.$$

As we shall see in a moment the hypothesis that the stress is zero in the initial state (characterised by $\epsilon_s^r = 0$) forces $\alpha = 0$ and the assumption of the infinitesimal theory is

$$\phi' = \rho_0\phi = \frac{\lambda + 2\mu}{2} I_1^2 - 2\mu I_2$$

(where, of course, the invariants I_1, I_2 of the classical theory are the approximations obtained by neglecting the second order terms in the expressions for the strain components). In the usual presentations of the infinitesimal theory the invariant $I'_2 = \epsilon_\beta^\alpha\epsilon_\alpha^\beta = I_1^2 - 2I_2$ is used instead of our I_2 and the expression for ϕ' appears as

$$\phi' = \frac{\lambda}{2} I_1^2 + \mu I'_2.$$

On using the relations

$$\partial I_1/\partial \epsilon_r{}^s = \delta_s{}^r; \qquad \partial I_2/\partial \epsilon_r{}^s = \delta_s{}^r I_1 - \epsilon_s{}^r \qquad \partial I_3/\partial \epsilon_r{}^s = I_3 \zeta_s{}^r$$

where $\zeta_s{}^r$ is the tensor reciprocal to $\epsilon_s{}^r$:

$$\zeta_a{}^r \epsilon_s{}^a = \delta_s{}^r$$

and remembering that $\rho/\rho_0 = \sqrt{1 - 2I_2 + 4I_2 - 8I_3} = 1 - I_1 + \cdots$ we find

$$T_s{}^r = (1 - I_1 + \cdots) \left\{ \begin{array}{l} (\alpha + \lambda I_1 + (3l + m) I_1^2 + m I_2 - 2n I_3) \delta_s{}^r \\ + [2(\mu - \alpha) - (m + 2\lambda) I_1 - 2(3l + m) I_1^2 - 2m I_2] \epsilon_s{}^r \\ + 2(m I_1 - 2\mu) \epsilon_a{}^r \epsilon_s{}^a + n I_3 \zeta_s{}^r \end{array} \right\}.$$

The value of $T_s{}^r$ in the initial position ($\epsilon_s{}^r = 0$) is, accordingly

$$(T_s{}^r)_0 = \alpha \delta_s{}^r$$

(so that the assumption of isotropy forces the stress in the initial position to be scalar i. e. of the nature of a hydrostatic pressure); we shall make the usual assumption that the stress in the initial position is zero forcing $\alpha = 0$. On multiplying out by the factor $1 - I_1 + \cdots$ and neglecting quantities of higher order than the second we find

$$\begin{aligned} T_s{}^r = \{ & I_1 + (3l + m - \lambda) I_1^2 + m I_2 \} \delta_s{}^r \\ & + \{ 2\mu - (m + 2\lambda + 2\mu) I_1 \} \epsilon_s{}^r - 4\mu \epsilon_a{}^r \epsilon_s{}^a + n I_3 \zeta_s{}^r. \end{aligned}$$

We shall not keep in the following the quantities arising from the third order terms in the expansion of $\rho_0 \phi$ i. e. we shall set $l = 0$, $m = 0$, $n = 0$ when we obtain

$$T_s{}^r = \lambda I_1 (1 - I_1) \delta_s{}^r + 2\{ \mu - (\lambda + \mu) I_1 \} \epsilon_s{}^r - 4\mu \epsilon_a{}^r \epsilon_s{}^a.$$

The first invariant $T = T_a{}^a$ of the stress tensor is, accordingly,

$$\begin{aligned} T &= 3\lambda I_1 (1 - I_1) + 2\{ \mu - (\lambda + \mu) I_1 \} I_1 - 4\mu I'_2 \\ &= (3\lambda + 2\mu) I_1 - (5\lambda + 6\mu) I_1^2 + 8\mu I_2. \end{aligned}$$

If we neglect the second order terms we obtain the expressions of the infinitesimal theory

$$T_s{}^r = \lambda I_1 \delta_s{}^r + 2\mu \epsilon_r{}^s; \qquad T = (3\lambda + 2\mu) I_1.$$

Although, as we shall see, a good approximation to the results of experiments may be secured by using only two elastic constants λ, μ i. e. by neglecting the

third order terms in the expansion of $\rho_0\phi$ a true second order approximation would keep these third order terms thus introducing five elastic constants λ, μ, l, m, n. Neglecting, however, third order terms in the expressions for the stress tensor the second order approximation is

$$\begin{aligned} T_s{}^r &= \{\lambda I_1 + (3l + m - \lambda)I_1{}^2 + mI_2\}\delta_s{}^r \\ &\quad + \{2\mu - (m + 2\lambda + 2\mu)I_1\}\epsilon_s{}^r - 4\mu\epsilon_a{}^r\epsilon_s{}^a + nI_3\zeta_s{}^r \\ T &= (3\lambda + 2\mu)I_1 + (9l + 2m - 5\lambda - 6\mu)I_1{}^2 + (3m + n + 8\mu)I_2 \end{aligned}$$

(the invariant $\zeta_a{}^a$ of the reciprocal of the strain tensor $\epsilon_s{}^r$ being I_2/I_3). In the treatment by Seth (**7**) of the problem of finite strain the strain components were not truncated, as in the classical infinitesimal theory, by the omission of the second order terms, but the equations

$$T_s{}^r = \lambda I_1\delta_s{}^r + 2\mu\epsilon_s{}^r; \qquad T = (3\lambda + 2\mu)I_1$$

of the infinitesimal theory were taken over with the following explanatory remark " Since this is the simplest tensor form that we can take, it is quite natural for us to assume that the stress-strain relations are governed by equations of the above type." From the discussion given above it is clear that simplicity is not a sufficiently compelling reason; for the whole strength obtained by a willingness to keep the second order terms in the strain components is sacrificed by the omission of second order terms—such as those that occur in the terms $-\lambda I_1{}^2\delta_s{}^r$ etc.,—in the expressions for the stress components.

4. The case of hydrostatic pressure; comparison of theory with experiment. In this simplest case the strain tensor is scalar $\epsilon_s{}^r = \epsilon\delta_s{}^r$ and $2\epsilon = 1 - (V_0/V)^{2/3}$. In most cases the stress is a pressure rather than a tension and ϵ is negative and so we write $T_s{}^r = -p\delta_s{}^r$; $\epsilon = -f$. Then $I_1 = -3f$; $I_2 = 3f^2$; $I_3 = -f^3$, and the formula connecting pressure with change of volume is

$$\begin{aligned} p &= af + bf^2; \qquad & f &= \tfrac{1}{2}\{(V_0/V)^{2/3} - 1\} \\ a &= 3\lambda + 2\mu; \qquad & b &= 15\lambda + 10\mu - 27l - 9m - n. \end{aligned}$$

Observe that if we avail ourselves only of the two elastic constants λ, μ of the classical (infinitesimal) theory $b = 15\lambda + 10\mu = 5a$ so that our formula is a one constant one

$$p = a(f + 5f^2); \qquad f = \tfrac{1}{2}\{(V_0/V)^{2/3} - 1\}; \qquad a = 3\lambda + 2\mu.$$

For small compressions

$$f = \tfrac{1}{2}\{(V_0/V)^{2/3} - 1\} = \tfrac{1}{2}\{(1 - \Delta V/V_0)^{-(2/3)} - 1\}$$
$$= \tfrac{1}{3}(\Delta V/V_0) + \cdots ; \qquad \Delta V = V_0 - V$$

and so the modulus of compression: $p \div \Delta V/V_0 = p \div 3f$; hence the modulus of compression is $a/3$.* The following table gives a comparison of the theory with the experimental results of Bridgman (**10**) upon the compressibility of sodium. The quadratic formula $p = af + bf^2$ was used and the two constants at our disposal were determined by the experimental results at 2,000 atm. (the beginning of the experiment) and at 12,000 atm. (near the middle of the experimental range which ran from 2,000 to 20,000 atm., at intervals of 2,000 atm.). In this way the availability of the formula for purposes of *extrapolation* (12,000 to 20,000 atm.), as well as interpolation (2,000 to 12,000 atm.), was tested.

TABLE 1.

p	$\Delta V/V_0$	f	p (calculated)
2000	.0295	.0101	——
4000	.0552	.0193	4005
6000	.0779	.0278	6022
8000	.0981	.0356	8005
10000	.1165	.0430	10003
12000	.1332	.0500	——
14000	.1488	.0567	14008
16000	.1632	.0631	16014
18000	.1767	.0692	18006
20000	.1894	.0751	20007

$$a = 1.874 \times 10^5; \qquad b = 1.052 \times 10^6.$$

The largest discrepancy is that corresponding to a measured pressure of 6,000 atm., and a calculated pressure of 6,022 atm., an error of about one-third of 1%. The other calculated values are correct to within one-tenth of 1%, most being much closer. Attention should be called to the fact that $b = 5.6\ a$, so that a neglect of the third order terms in the expansion of $\rho_0\phi$ (which neglect would force b to be $5a$) would only disturb the agreement to the extent $.6af^2$. Thus the one constant formula $p = a(f + 5f^2)$ fits the data for sodium to an accuracy of within 1.5% over the range 2,000 atm., to 20,000 atm., the constant a having the value 1.92×10^5 (determined by the measurement at 12,000 atm.). The agreement is as good as could be expected since f is only

* again for small compressions

① Bulk modulus?

measured to an accuracy varying between .5% at 2,000 atm., to .07% at 20,000 atm.

A two constant formula does not give very good agreement with experimental results for liquids which are much more compressible than solids. However a three constant formula (which would result if the energy-density were expanded as far as fourth order terms) gives very good agreement. The following table gives the result of a comparison between the results of calculation from a three-constant formula $p = af + bf^2 + cf^3$ and the experimental results of Bridgman (**11**) on the compressibility of N-Amyl Iodide, at 0° temperature, under pressures varying between 500 atm., and 12,000 atm. Over this range the pressures calculated agreed with those measured to within less than one per cent.

TABLE 2.

N-Amyl Iodide (0°)

p (obs)	V/V_0	f	af	bf^2	cf^3	p (calc.)
500	.9685	.0108	444.9	49.4	10.2	——
1000	.9442	.0195	803.9	161.7	37.2	1002.8
1500	.9250	.0267	1099	302.3	134.7	1496.4
2000	.9094	.0327	1347	453.7	174.9	1975.6
3000	.8831	.0432	1780	792.9	404.2	2977.1
4000	.8624	.0518	2136	1142	698.7	3976.7
5000	.8451	.0593	2445	1496	1048	4989
6000	.8304	.0659	2717	1848	1438	——
7000	.8173	.0720	2967	2202	1871	7040
8000	.8064	.0771	3177	2525	2297	7999
9000	.7965	.0819	3375	2849	2753	8977
10000	.7873	.0864	3560	3171	3233	9964
11000	.7786	.0908	3741	3502	3752	10995
12000	.7706	.0948	3908	3822	4277	——

$$a = 512.04 \times 10^2; \qquad b = 424.8 \times 10^3; \qquad c = 501.2 \times 10^4.$$

In order to show the dependence of the coefficients a, b, c upon the temperature similar calculations were made from measurements at 50° C with the following result

$$a_{50} = 303.94 \times 10^2; \qquad b_{50} = 316.26 \times 10^3; \qquad c_{50} = 365.51 \times 10^4.$$

A similar three constant formula for sodium over the range 2,000 to 20,000 atmospheres (the constants being determined by the measurements at 2,000, 10,000 and 20,000 atmospheres) gave the values

$$a = 188.13 \times 10^3; \quad b = 101.1 \times 10^4; \quad c = 37.2 \times 10^4$$

and the correspondence between the observed and calculated values was:

p (obs)	4000,	6000,	8000,	12000,	14000,	16000,	18000
p (calc)	4008,	6014,	8003,	11979,	13984,	15976,	17984

an agreement to within one quarter of one per cent over the entire range.

5. The Young's modulus experiment. In this experiment the ends of a cylinder of length l are subjected to a uniform stress T, the sides being free from any applied force. The conditions of the problem are met by assuming $u = px$; $v = py$; $w = rz$ (where the z-axis is parallel to the generators of the cylinder) it being agreed that mass forces, such as the weight of the mass elements of the cylinder, may be neglected. The strain tensor is diagonal with diagonal elements

$$\epsilon_1{}^1 = \epsilon_2{}^2 = p - \tfrac{1}{2}p^2; \qquad \epsilon_3{}^3 = r - \tfrac{1}{2}r^2$$

and the stress tensor is consequently also diagonal. The numbers p, r must be such that $T_x{}^x = T_y{}^y = 0$; $T_z{}^z = T$. If $e = w/c$ is the relative extension we have, since $z = c + w$, $e = r(1 + e)$ so that $r = e - e^2 + \cdots$ implying

$$\epsilon_3{}^3 = e - (3/2)e^2 + \cdots; \qquad e = \epsilon_3{}^3 + (3/2)(\epsilon_3{}^3)^2 + \cdots.$$

For simplicity of notation we write ϵ for $\epsilon_3{}^3$ and set $\epsilon_1{}^1 = \epsilon_2{}^2 = -\sigma\epsilon_3{}^3$. If $f = -u/a = -v/b$ denotes, for the moment, the relative contraction, in a direction perpendicular to the applied stress, we have

$$\epsilon_1{}^1 = -f - \frac{3}{2}f^2 + \cdots \text{ so that } \sigma = (f + \frac{3}{2}f^2 + \cdots)/(e - \frac{3}{2}e^2 + \cdots)$$
$$= \frac{f}{e}\{1 + \frac{3}{2}(f + e) + \cdots\}.$$

To a first approximation $\sigma = f/e$ measures the ratio of the relative contraction, perpendicular to the applied stress, to the relative extension in the direction of the applied stress; we shall refer to σ as Poisson's ratio.

Since $T_x{}^x = T_y{}^y = 0$ we have $T = T_z{}^z = T_\alpha{}^\alpha$ and so

$$T = (3\lambda + 2\mu)I_1 + (9l + 2m - 5\lambda - 6\mu)I_1{}^2 + (3m + n + 8\mu)I_2.$$

We shall content ourselves with examining what partial explanation of the phenomena observed in the Young's modulus experiment may be obtained by

using only the two elastic constants λ, μ of the infinitesimal theory. On setting $l=0$, $m=0$, $n=0$ we find

$$T=(3\lambda+2\mu)I_1-(5\lambda+6\mu)I_1^2+8\mu I_2.$$

On putting in the values $I_1=(1-2\sigma)\epsilon$, $I_2=(\sigma^2-2\sigma)\epsilon^2$ we find

$$T=(3\lambda+2\mu)(1-2\sigma)\epsilon-\{(5\lambda+6\mu)(1-2\sigma)^2+8\mu(2\sigma-\sigma^2)\}\epsilon^2.$$

The relation $T_x{}^x=0$ yields

$$\lambda-2\sigma(\lambda+\mu)-\{8(\lambda+\mu)\sigma^2-2(3\lambda+\mu)\sigma+\lambda\}\epsilon=0$$

or

$$\left\{\sigma-\frac{\lambda}{2(\lambda+\mu)}\right\}\{1+(4\sigma-1)\epsilon\}=0.$$

Since $V_0/V=(1+2\sigma\epsilon)\sqrt{1-2\epsilon}$ the maximum value of ϵ is .5 and so, granting $\sigma>0$, $1+(4\sigma-1)\epsilon>.5$ so that $\sigma=\lambda/2(\lambda+\mu)$. It is important to notice that this constancy of Poisson's ratio σ is not a mere approximation but an exact result. On inserting this value of σ in the expression for T we find

$$T=E\epsilon\left\{1-\frac{2\lambda+3\mu}{\lambda+\mu}\epsilon\right\}$$

where $E=\dfrac{(3\lambda+2\mu)\mu}{\lambda+\mu}$ is Young's modulus. The first approximation, which would be furnished by the infinitesimal theory, is $T=E\epsilon$ but the significance of the quadratic formula of the finite theory is that the graph of T against ϵ is a parabola instead of a straight line. Hence T has a maximum value $(T)_{\max}=\dfrac{\lambda+\mu}{4(2\lambda+3\mu)}E$ occurring when $\epsilon=\dfrac{\lambda+\mu}{2(2\lambda+3\mu)}$. What this means is that if a larger stress is applied the deformation cannot be of the simple type described by the formulae $u=px$; $v=py$; $w=rz$. For steel λ is approximately 1.5μ so that $(T)_{\max}=E/10$ the corresponding value of ϵ being .2. In using the formula given above for T it should be noticed that $\epsilon=r-\frac{1}{2}r^2=\dfrac{(2+e)e}{2(1+e)^2}$ where e is the relative extension. These results of the finite theory, predicting a maximum stress (yield point) are qualitatively correct but the predicted value of $(T)_{\max}$ is much too large.

6. The stress-strain equations for a non-isotropic medium. The elastic potential is now a function $\phi({}_{pq}\eta, {}_{pq}c, T)$ of the scalar strain components ${}_{pq}\eta$, the coefficients ${}_{pq}c$ of the quadratic form for $(ds_0)^2$ and the temperature T. On using the expression for $\delta\,{}_{pq}\eta$ given in 1, namely

$$2\delta\, {}_{pq}\eta = \{(\delta x_a)_{,\beta} + (\delta x_\beta)_{,a}\}({}_{p,}x^a)({}_{q,}x^\beta)$$

and agreeing that ϕ is symmetrized with respect to the scalar strain components ${}_{pq}\eta$ so that $\partial\phi/\partial({}_{pq}\eta) = \partial\phi/\partial({}_{qp}\eta)$ we find at once

$$T^{rs} = \rho \frac{\partial\phi}{\partial({}_{a\beta}\eta)}\, {}_{a,}x^r\, {}_{\beta,}x^s.$$

These expressions give, from the Lagrangian viewpoint, the stress tensor in terms of the gradient of the elastic potential relative to the scalar strain components ${}_{pq}\eta$. In order to obtain the corresponding Eulerian equations we introduce the matrix $\boldsymbol{j}$ reciprocal to the matrix $\boldsymbol{k}$ whose elements

$${}_{pq}k = g_{a\beta}({}_{p,}x^a)({}_{q,}x^\beta) = 2\, {}_{pq}\eta + {}_{pq}c\,;$$

it being clear that the elements of $\boldsymbol{j}$ are given by the formula

$${}^{pq}j = g^{a\beta}({}^{p}a_{,a})({}^{q}a_{,\beta}).$$

On taking the variation of the matrix equation $\boldsymbol{jk} = \boldsymbol{e}$ (the unit matrix) we find $\delta\boldsymbol{j}\cdot\boldsymbol{k} + \boldsymbol{j}\cdot\delta\boldsymbol{k} = 0$ or equivalently, $\delta\boldsymbol{j} = -\boldsymbol{j}\cdot\delta\boldsymbol{k}\cdot\boldsymbol{j}$. Hence

$$\frac{\partial\phi}{\partial({}_{a\beta}\eta)}\delta({}_{a\beta}\eta) = \frac{\partial\phi}{\partial({}^{a\beta}j)}\delta({}^{a\beta}j) = -\frac{\partial\phi}{\partial({}^{a\beta}j)}({}^{a\sigma}j)\delta({}_{\sigma\tau}k)({}^{\tau\beta}j)$$
$$= -2\frac{\partial\phi}{\partial({}^{a\beta}j)}({}^{a\sigma}j)\delta({}_{\sigma\tau}\eta)({}^{\tau\beta}j)$$

so that

$$\frac{\partial\phi}{\partial({}_{pq}\eta)} = -2({}^{pa}j)\frac{\partial\phi}{\partial({}^{a\beta}j)}({}^{\beta q}j).$$

Since

$$({}^{pa}j)({}_{a,}x^r) = g^{ra}({}^{p}a_{,a}) = {}^{p}a^{,r}$$

it follows that

$$T^{rs} = -2\rho\frac{\partial\phi}{\partial({}^{a\beta}j)}({}^{a}a^{,r})({}^{\beta}a^{,s}).$$

Appendix.

Expression of the quantities J_1, J_2, J_3 in terms of the invariants I_1, I_2, I_3. The invariants s_1, s_2, s_3 of any matrix $\boldsymbol{u}$ are determined by the equation

$$\det(\lambda\delta_s{}^r - u_s{}^r) = \lambda^3 - s_1\lambda^2 + s_2\lambda - s_3.$$

It is immediately clear that if $\boldsymbol{u}, \boldsymbol{v}$ are any two non-singular 3×3 matrices the two products $\boldsymbol{uv}$ and $\boldsymbol{vu}$ have the same invariants. In fact if $\boldsymbol{E} = (\delta_s{}^r)$ is the three-rowed unit matrix

$$\det(\lambda \boldsymbol{E} - \boldsymbol{vu}) = \det \boldsymbol{u}^{-1}(\lambda \boldsymbol{E} - \boldsymbol{uv})\boldsymbol{u} = \det(\lambda \boldsymbol{E} - \boldsymbol{uv}).$$

Secondly if $\boldsymbol{u}$ is any non-singular three-rowed matrix with invariants s_1, s_2, s_3 the invariants $\sigma_1, \sigma_2, \sigma_3$ of its reciprocal $\boldsymbol{u}^{-1}$ are given by the formulae $\sigma_1 = s_2/s_3$; $\sigma_2 = s_1/s_3$; $\sigma_3 = 1/s_3$. For

$$\det(\lambda \boldsymbol{E} - \boldsymbol{u}^{-1}) = -\lambda^3 (\det \boldsymbol{u})^{-1} \det\left(\frac{1}{\lambda}\boldsymbol{E} - \boldsymbol{u}\right) = \frac{1}{s_3}\{\lambda^3 s_3 - \lambda^2 s_2 + \lambda s_1 - 1\}$$

since $\det \boldsymbol{u} = s_3$.

Now if we denote by $\boldsymbol{g}$ the matrix whose elements are g_{pq} and by $\boldsymbol{t}$ the matrix whose elements are ${}_{s,}x^r$ (the upper label denoting the row and the lower the column) the matrix whose elements are ${}_{pq}k = g_{\alpha\beta}({}_{p,}x^\alpha)({}_{q,}x^\beta)$ is $\boldsymbol{t'gt}$ where the prime attached to a matrix denotes, as usual, its transposed. Hence the matrix $\boldsymbol{k}$ whose elements are

$${}^p{}_q k = ({}^{p\alpha}c)({}_{\alpha q}k) = 2({}^p{}_q\eta) + {}^p{}_q\delta$$

is $\boldsymbol{c}^{-1}\boldsymbol{t'gt}$ ($\boldsymbol{c}$ being the matrix whose elements are ${}_{pq}c$). Similarly the matrix whose elements are $h_{pq} = {}_{\alpha\beta}c({}^\alpha a_{,p})({}^\beta a_{,q})$ is $\boldsymbol{t}^{-1\prime}\boldsymbol{ct}^{-1}$ since the matrix whose elements are ${}^r a_{,s}$ is the reciprocal of the matrix whose elements are ${}_{s,}x^r$. Hence the matrix $\boldsymbol{h}$ whose elements are $h_q{}^p = \delta_q{}^p - 2\epsilon_q{}^p$ is $\boldsymbol{g}^{-1}(\boldsymbol{t}^{-1})'\boldsymbol{ct}^{-1}$. Hence its reciprocal $\boldsymbol{h}^{-1}$ is given by $\boldsymbol{h}^{-1} = \boldsymbol{tc}^{-1}\boldsymbol{t'g}$ and since $\boldsymbol{k} = \boldsymbol{c}^{-1}\boldsymbol{t'gt}$ it follows that $\boldsymbol{k}$ and $\boldsymbol{h}^{-1}$ have the same invariants. But the invariants of $\boldsymbol{k}$ are the coefficients of $-\lambda^2, \lambda, -1$, respectively, in the development of

$$\begin{aligned}\det(\lambda \boldsymbol{E} - \boldsymbol{k}) &= \det[(\lambda - 1)\boldsymbol{E} - 2\boldsymbol{\eta}] \\ &= 8\det(\mu \boldsymbol{E} - \boldsymbol{\eta}); \qquad \mu = \frac{\lambda - 1}{2} \\ &= (\lambda - 1)^3 - 2(\lambda - 1)^2 J_1 + 4(\lambda - 1)J_2 - 8J_3.\end{aligned}$$

Hence the invariants of $\boldsymbol{k}$ are, respectively,

$$2J_1 + 3, \qquad 4J_2 + 4J_1 + 3, \qquad 8J_3 + 4J_2 + 2J_1 + 1.$$

Similarly the invariants of $\boldsymbol{h}$ are, respectively,

$$3 - 2I_1, \qquad 3 - 4I_1 + 4I_2, \qquad 1 - 2I_1 + 4I_2 - 8I_3$$

so that the invariants of $\boldsymbol{h}^{-1}$ are, respectively,

$$\frac{3-4I_1+4I_2}{1-2I_1+4I_2-8I_3}; \qquad \frac{3-2I_1}{1-2I_1+4I_2-8I_3}; \qquad \frac{1}{1-2I_1+4I_2-8I_3}.$$

Hence we have the equations

$$2J_1+3=\frac{3-4I_1+4I_2}{1-2I_1+4I_2-8I_3}; \quad 4J_2+4J_1+3=\frac{3-2I_1}{1-2I_1+4I_2-8I_3},$$
$$8J_3+4J_2+2J_1+1=\frac{1}{1-2I_1+4I_2-8I_3}$$

and solving these we find

$$J_1=\frac{I_1-4I_2+12I_3}{1-2I_1+4I_2-8I_3}; \; J_2=\frac{I_2-6I_3}{1-2I_1+4I_2-8I_3}; \; J_3=\frac{I_3}{1-2I_1+4I_2-8I_3}$$

or, equivalently,

$$I_1=\frac{J_1+4J_2+12J_3}{1+2J_1+4J_2+8J_3}; \; I_2=\frac{J_2+6J_3}{1+2J_1+4J_2+8J_3}; \; I_3=\frac{J_3}{1+2J_1+4J_2+8J_3}.$$

Summary.

In the present paper formulae are derived which enable one to calculate the stress in an elastic medium when the strain and the elastic energy density are known, no simplifying assumptions, such as smallness of strain, being necessary. For an isotropic elastic solid under hydrostatic pressure the following *one* constant formula gives good agreement with experimental observation (only two elastic constants λ, μ being used in the expression for the elastic energy density)

$$p=a(f+5f^2); \qquad f=\tfrac{1}{2}\{(V_0/V)^{2/3}-1\}; \qquad a=3\lambda+2\mu.$$

In the Young's modulus experiment the formula for the extensional stress (again using only the two constants λ, μ) is

$$T=E\epsilon\left\{1-\frac{2\lambda+3\mu}{\lambda+\mu}\epsilon\right\}; \qquad E \text{ (Young's modulus)}=\frac{(3\lambda+2\mu)\mu}{\lambda+\mu}$$

where $\epsilon=\frac{(2+e)e}{(1+e)^2}$, e being the relative extension. Hence T has a maximum value $\frac{\lambda+\mu}{4(2\lambda+3\mu)}E$, occurring when $\epsilon=\frac{\lambda+\mu}{2(2\lambda+3\mu)}$. For a true second

order approximation (the infinitesimal theory being regarded as a first order approximation) five elastic constants occur and the corresponding formulae are either given or their derivation is immediate.

THE JOHNS HOPKINS UNIVERSITY;
INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

1. G. Kirchhoff, "Über die Gleichungen des Gleichgewichtes eines elastischen Körpers bei nicht unendlich kleinen Verschiebungen seiner Theile," *Sitz. math-nat. Klasse der kaiserlichen Akad. der Wiss.*, **9** (1852), p. 762.
2. J. Boussinesq, "Théorie des ondes liquides périodiques," *Mémoires présentés à l'Ac. des Sciences*, **20** (1869), p. 516.
3. E. et F. Cosserat, "Sur la théorie de l'élasticité," *Ann. de Toulouse*, **10** (1896).
4. P. Duhem, "Recherches sur l'Élasticité," *Ann. de l'École nor. sup.*, 3me Série, **21**, **22**, **23** (1904-06).
5. E. Almansi, "Sulle deformazioni finite dei solidi elastici isotropi," *Rend. Lincei*, Ser. 5, Tome **20**[1] (1911).
6. L. Brillouin, "Sur les tensions de radiation," *Ann. de Physique*, 10me Série **4** (1925).
7. B. R. Seth, "Finite strain in elastic problems," *Phil. Trans. Roy. Soc.*, A **234** (1935), p. 231.
8. A. Signorini, "Trasformazioni termoelastiche finite, etc.," *Soc. Italiana per il progresso delle scienze* **14** (1936).
9. E. G. Coker, and L. N. G. Filon, *Treatise on photo-elasticity*, Cambridge (1931).
10. P. W. Bridgman, "Electrical resistances and volume changes up to 20,000Kg./cm.^{2}," *Proceedings of the National Academy of Sciences*, **21** (1935), p. 109.
11. P. W. Bridgman, "The pressure-volume-temperature relations of fifteen liquids," *Proc. Amer. Ac. Arts and Sc.*, **68** (1933).

MEAN MOTIONS AND DISTRIBUTION FUNCTIONS.[1]

By Philip Hartman, E. R. van Kampen and Aurel Wintner.

Let $k = 1, \cdots, n$ and $\Sigma = \sum_{k=1}^{n}$. If $z(t)$ is a function of the form

$$(1) \qquad z(t) \equiv x(t) + iy(t) = \Sigma\, a_k \exp 2\pi i(\lambda_k t + \alpha_k),$$

where λ_k, α_k are real and $a_k > 0$, put

$$(2) \qquad z(t) = \pm \mid z(t) \mid \exp 2\pi i \phi(t),$$

where the sign of $\pm \mid z(t) \mid$ is to be chosen for every t in such a way that $\phi = \phi(t)$ becomes a continuous function of t. The function $\phi(t)$ is said to have for $t \to +\infty$ a mean motion μ if

$$(3) \qquad \phi(t)/t \to \mu, \text{ i. e., } \phi(t) = \mu t + o(t); \; t \to +\infty.$$

The problem of the existence and the determination of this constant μ goes back to Lagrange's approximative treatment of secular perturbations of the major planets and has been solved in the case $n = 3$ by Bohl.[2] The case $n = 4$ has been treated by Weyl.[3] The present note attempts a general approach to the problem from the point of view of the theory of distribution functions.

The connection between mean motions and asymptotic distribution functions is suggested by the following consideration: It is known [4] that if the real function $\xi = \psi(t)$ is almost periodic in the sense of Bohr, then $\psi(t)$ possesses an asymptotic distribution function $\sigma(\xi)$ and one has, for every continuous function $f = f(\xi)$,

$$(4) \qquad \int_{-\infty}^{+\infty} f(\xi)\, d\sigma(\xi) = M\{f(\psi(t))\},$$

where

$$(5) \qquad M\{g(t)\} = \lim_{T \to +\infty} (1/T) \int_0^T g(t)\, dt.$$

It is understood that by the existence of an asymptotic distribution function of a real measurable function $\psi(t)$, where $0 \leqq t < +\infty$, is meant the existence of a monotone function $\sigma(\xi)$ such that $\sigma(-\infty) = 0$, $\sigma(+\infty) = 1$ and, if ξ is a continuity point of $\sigma(\xi)$,

[1] Received January 20, 1937.

[2] Bohl [3].

[3] Weyl [7].

[4] Wintner [8].

$$(1/T) \text{ meas } [\psi(t) \leqq \xi]_T \to \sigma(\xi), \qquad T \to +\infty,$$

where $[\psi(t) \leqq \xi]_T$ denotes the set of those points t of the interval $0 \leqq t \leqq T$ at which $\psi(t) \leqq \xi$. Now suppose that the amplitudes a_k of (1) satisfy the condition of Lagrange, i. e., that

$$a_{k_1} > a_{k_2} + \cdots + a_{k_n} \tag{6}$$

holds for a permutation $(k_1, \cdots, k_n)$ of $(1, \cdots, n)$. Then (3) may be replaced by the sharper statement that

$$\phi(t) = \mu t + \omega(t), \tag{3 bis}$$

where $\omega(t)$ and also its derivative $\omega'(t)$ are almost periodic in the sense of Bohr.[5] Hence, on denoting by $\psi(t)$ the almost periodic function $\phi'(t) = \mu + \omega'(t)$ and by $\sigma(\xi)$ its asymptotic distribution function, (4) is applicable, and becomes, when $f(\xi) = \xi$,

$$\int_{-\infty}^{+\infty} \xi d\sigma(\xi) = M\{\phi'(t)\}, \tag{7}$$

where, according to (5) and (3),

$$M\{\phi'(t)\} = \lim_{T \to +\infty} \phi(T)/T = \mu. \tag{8}$$

Thus

$$\int_{-\infty}^{+\infty} \xi d\sigma(\xi) = \mu, \tag{9}$$

so that the mean motion of $\phi(t)$ appears as the first moment of the asymptotic distribution function $\sigma(\xi)$ of $\phi'(t)$.

If the amplitudes a_k do not satisfy the inequality (6), the preceding considerations break down in view of the fact that $\phi'(t)$ is not, in general, almost periodic in the sense of Bohr, while a possible treatment of the problem within a class of almost periodic functions more general than those of Bohr leads to difficulties. In fact, one needs the particular case (7) of (4), and (7) is obvious only when ψ $(= \phi')$ is a bounded function, a condition which is not satisfied in the majority of cases, if (6) does not hold. It is not difficult to prove that the function $\phi'(t)$ has a distribution function $\sigma(\xi)$ and that the space-average represented by the Stieltjes integral on the left of (7) is absolutely convergent. The main difficulty arises in the identification of the

[5] Wintner [9].

space-average with the corresponding time-average $M\{\phi'(t)\}$. It is known [6] that the truth of Lindelöf's hypothesis in the theory of the Riemann zeta-function depends on a question of the same type, namely on the question as to the admissibility of the identification of certain space-averages with the corresponding (hypothetical) time-averages, as expressed by (4).

It will be assumed that the frequencies λ_k of (1) are linearly independent. This assumption, as will be seen from the proof, does not essentially affect the validity of the method. After proving the existence of the asymptotic distribution function $\sigma(\xi)$ and of its first moment, the identification of the time-average with the space-average remains to be treated. The admissibility of this identification will be proved with the help of Birkhoff's ergodic theorem,[7] by excluding, for fixed values of the amplitudes a_k and the frequencies λ_k, a set of measure zero in the n-dimensional space of the phases α_k. Actually, there are some indications that Birkhoff's zero set is empty in the present case. The assumption of the linear independence of the frequencies λ_k is to the effect [8] that the problem is of the metrically transitive type, so that the mean motion μ will depend on the amplitudes a_k and the frequencies λ_k but not on the phases α_k.

The explicit evaluation of μ will be reduced for every n to the evaluation of a definite integral in the α_k-space. On comparing the results of the present paper with those of Bohl [2] $(n=3)$, it follows that if a_1, a_2, a_3 are positive, then the integral

$$\int_0^1\int_0^1\int_0^1 \frac{[a_1^2 + a_1a_2\cos 2\pi(\vartheta_2-\vartheta_1) + a_1a_3\cos 2\pi(\vartheta_3-\vartheta_1)]\,d\vartheta_1\,d\vartheta_2\,d\vartheta_3}{a_1^2+a_2^2+a_3^2+2a_1a_2\cos 2\pi(\vartheta_2-\vartheta_1)+2a_1a_3\cos 2\pi(\vartheta_3-\vartheta_1)+2a_2a_3\cos 2\pi(\vartheta_3-\vartheta_2)}$$

$$=\int_0^1\int_0^1 \frac{[a_1^2 + a_1a_2\cos 2\pi\vartheta_1 + a_1a_3\cos 2\pi\vartheta_2]\,d\vartheta_1\,d\vartheta_2}{a_1^2+a_2^2+a_3^2+2a_1a_2\cos 2\pi\vartheta_1+2a_1a_3\cos 2\pi\vartheta_2+2a_2a_3\cos 2\pi(\vartheta_1-\vartheta_2)}$$

is equal to

$$\frac{1}{\pi}\operatorname{arc\,cos}(a_2^2+a_3^2-a_1^2)/(2a_2a_3)$$

if $a_i \leqq a_j + a_k$ for all permutations (i, j, k) of $(1, 2, 3)$, while it is equal to 1 if $a_1 > a_2 + a_3$, and finally it is equal to 0 if either $a_2 > a_1 + a_3$ or $a_3 > a_1 + a_2$. Similar relations follow by comparison of the results of the present paper with those of Weyl [3] $(n=4)$. These definite integrals show with varying a_k "discontinuities" of the same type as the well-known "discontinuous" integrals

[6] Cf., on the one hand, Hardy and Littlewood [4] and, on the other hand, Jessen and Wintner [5], Theorem 31.

[7] Birkhoff [1]; cf. also Khintchine [6].

[8] Cf. Birkhoff [2], p. 371.

of Sonine ($n=3$) and Nicholson ($n=4$), occurring in connection with Lord Rayleigh's random walk problem and presented on pp. 411, 414 and 420 of Watson's Treatise on Bessel functions.

For a given function (1), put

$$(10)\qquad \psi(t)=\tfrac{1}{2\pi}\,\frac{x(t)y'(t)-y(t)x'(t)}{[x(t)]^2+[y(t)]^2},\ \text{if } z(t)\neq 0,$$

where the prime denotes differentiation with respect to t. It is clear that, no matter which locally continuous determination is chosen for

$$\arg z(t)=-i\log[z(t)/|z(t)|],$$

one has

$$[\arg z(t)]'=2\pi\psi(t),\ \text{if } z(t)\neq 0.$$

It is known [2] that a unique function $\phi(t)$ is defined by the following requirements:

(i) $\phi(t)$ is continuous for $-\infty<t<+\infty$;
(ii) $\phi'(t)=\psi(t)$, if $z(t)\neq 0$;
(iii) $2\pi\phi(t)\equiv\arg z(t)\pmod{\pi}$, if $z(t)\neq 0$;
(iv) $0\leqq\phi(0)<\frac{1}{2}$.

The function $\phi(t)$ thus defined satisfies (2), where one has to choose, for every t, the sign $+$ or the sign $-$ according as the difference

$$|2\pi\phi(t)-\arg z(t)|$$

is an even or an odd multiple of π, while there is no ambiguity in the case $z(t)=0$. The function $\phi(t)$ will be referred to as the angular function belonging to (1). The values of t for which $z(t)=0$, i.e., for which $\phi'(t)=\psi(t)$ is undefined, clearly do not have a cluster point.

Let $\chi=\chi(\vartheta_1,\cdots,\vartheta_n)$ denote the function

$$(11)\quad \chi=\frac{(\Sigma\lambda_k a_k\cos 2\pi\vartheta_k)(\Sigma a_k\cos 2\pi\vartheta_k)+(\Sigma\lambda_k a_k\sin 2\pi\vartheta_k)(\Sigma a_k\sin 2\pi\vartheta_k)}{(\Sigma a_k\cos 2\pi\vartheta_k)^2+(\Sigma a_k\sin 2\pi\vartheta_k)^2}$$

defined on the n-dimensional torus

$$(12)\qquad \Theta:\ 0\leqq\vartheta_k<1;\qquad (k=1,\cdots,n),$$

except on the set N of those points of Θ at which the denominator of (11) vanishes, so that N is the set on Θ defined by the pair of equations

$$(13)\qquad \mathrm{N}:\ F\equiv\Sigma a_k\cos 2\pi\vartheta_k=0,\qquad G\equiv\Sigma a_k\sin 2\pi\vartheta_k=0.$$

Suppose that the frequencies λ_k of (1) are linearly independent and let

Z denote the curve on the torus (12) which is defined by the parameter representation

$$(14) \qquad Z: \quad \vartheta_k = \lambda_k t + \alpha_k \pmod 1; \qquad (k = 1, \cdots, n),$$

where the parameter t runs from 0 to $+\infty$ and the phases α_k are arbitrarily fixed.

First, it will be shown that the derivative $\phi'(t)$ of the angular function $\phi(t)$ of the function (1) has an asymptotic distribution function $\sigma(\xi)$ and that

$$(15) \qquad \sigma(\xi) = \text{meas}\, \Gamma_\xi, \qquad -\infty < \xi < +\infty,$$

where the meas Γ_ξ denotes the n-dimensional $(\vartheta_1, \cdots, \vartheta_n)$-measure of the set Γ_ξ of those points $(\vartheta_1, \cdots, \vartheta_n)$ of the torus (12) at which the function (11) satisfies the inequality

$$(16) \qquad \chi \leqq \xi.$$

In order to prove this, notice first that the set (13) at which the denominator of (11) vanishes clearly has a vanishing Jordan content (a detailed description of N will be given below). It is also clear from (11) that Γ_ξ has for every ξ a Jordan content. On the other hand, the linear independence of the λ_k implies that to distinct values of t there belong distinct points of the curve (14) on the torus (12). Furthermore, it is seen from (1), (10) and (11) that

$$(17) \qquad \chi(\lambda_1 t + \alpha_1, \cdots, \lambda_n t + \alpha_n) = \psi(t) = \phi'(t)$$

holds at those points of the curve (14) which do not lie on the subset (13) of the torus (12), i.e., at which $z(t) \neq 0$. Now it is clear from (17) and from the definition of Γ_ξ that $\psi(t) \leqq \xi$ holds if and only if that point of the curve (14) to which t belongs is a point of Γ_ξ. It follows, therefore, from the Kronecker-Weyl approximation theorem that if $\{\xi; T\}$ denotes the sum of the lengths of those subintervals of the interval $0 \leqq t \leqq T$ on which $\psi(t) \leqq \xi$, then

$$\{\xi; T\}/T \to \text{meas}\, \Gamma_\xi, \qquad T \to +\infty.$$

This proves that $\psi(t) = \phi'(t)$ has the function (15) as asymptotic distribution function.

Next, it will be shown that the set N defined by (13) is a closed, possibly disconnected, $(n-2)$-dimensional analytic manifold in the n-dimensional torus Θ defined by (12), and that the manifold N has no singularities or a finite number of singular curves according as there does not or does exist at least one permutation $(i_1, \cdots, i_n)$ of $(1, \cdots, n)$ such that

$$(18) \qquad a_{i_1} + \cdots + a_{i_m} = a_{i_{m+1}} + \cdots + a_{i_n}$$

holds for some m. It will be seen from the proof that if $n = 2$ or $n = 3$, then $\mathbf{N}$ consists of at most two analytic simple closed curves without singularities. Incidentally, $\mathbf{N}$ is, for an arbitrary n, empty if and only if the condition (6) of Lagrange is satisfied. This is clear from the definition (13) of $\mathbf{N}$.

In order to prove that $\mathbf{N}$ has the structure described above, let j, l be a pair of distinct values of $k = 1, \cdots, n$ and let J_{jl} denote the Jacobian with respect to ϑ_j, ϑ_l of the two functions F, G occurring in the definition (13) of $\mathbf{N}$, so that

$$(19) \qquad J_{jl} = \frac{\partial(F, G)}{\partial(\vartheta_j, \vartheta_l)} = 4\pi^2 a_j a_l \sin 2\pi(\vartheta_l - \vartheta_j); \qquad j \neq l.$$

Accordingly, a point $\mathbf{P} = (\vartheta_1, \cdots, \vartheta_n)$ of the subset $\mathbf{N}$ of the torus Θ is a singular point of the manifold $\mathbf{N}$ if and only if

$$(20) \qquad 2 \mid \vartheta_j - \vartheta_l \mid \equiv 0 \pmod 1), \quad \text{where} \quad j, l = 1, \cdots, n; \; j \neq l.$$

Suppose that there exists on $\mathbf{N}$ a point $\mathbf{P} = (\vartheta_1, \cdots, \vartheta_n)$ satisfying (20). Then one can arrange the numbers $\vartheta_1, \cdots, \vartheta_n$ into two groups

$$(\vartheta_{i_1}, \cdots, \vartheta_{i_m}), \qquad (\vartheta_{i_{m+1}}, \cdots, \vartheta_{i_n})$$

such that

$$(21a) \qquad \vartheta_{i_r} - \vartheta_{i_s} = \tfrac{1}{2}; \qquad (r = 1, \cdots, m; \; s = m + 1, \cdots, n),$$

while

$$(21b) \qquad \vartheta_{i_1} = \cdots = \vartheta_{i_m}, \qquad \vartheta_{i_{m+1}} = \cdots = \vartheta_{i_n}.$$

Now it is clear that the pair of conditions (21a), (21b) defines on the torus Θ a simple closed curve, and that this simple closed curve lies, in view of (20) and (13), on the $(n-2)$-dimensional manifold $\mathbf{N}$ if and only if the amplitudes a_k satisfy (18).

In what follows, it will be assumed for the sake of brevity that the amplitudes a_k of (1) do not satisfy an equation of the form (18), i. e., that the manifold $\mathbf{N}$ is free of singularities, so that no point of Θ which satisfies (20) is a point of $\mathbf{N}$.

In order to prove that the space-average occurring on the left-hand side of (7) is finite for the distribution function (15), one has, according to the definition of Γ_ξ, merely to prove that the function (11) is absolutely integrable over the torus Θ. Actually,

$$(22) \qquad \int\limits_{\Theta} |\chi|^\kappa \, d\Theta = \int_0^1 \cdots \int_0^1 |\chi(\vartheta_1, \cdots, \vartheta_n)|^\kappa \, d\vartheta_1 \cdots d\vartheta_n < +\infty, \text{ if } 0 < \kappa < 2.$$

The proof of (22) proceeds as follows: Let

$$(23) \qquad \mathbf{P}^0\colon \; \vartheta_k = \vartheta_k^0; \qquad (k = 1, \cdots, n),$$

be a point of the manifold N, so that, since N is free of singularities, (20) is not satisfied by $(\vartheta_1, \cdots, \vartheta_n) = (\vartheta_1^0, \cdots, \vartheta_n^0)$, and so there exists at least one pair j, l, say $j = 1$ and $l = 2$, such that

(24) $$2 \mid \vartheta_1^0 - \vartheta_2^0 \mid \not\equiv (\text{mod } 1), \text{ i. e., } \cos 2\pi(\vartheta_1^0 - \vartheta_2^0) \neq \pm 1.$$

Now the quadratic form

(25) $$Q(u, v; \vartheta_1, \vartheta_2) = 4\pi^2[a_1^2u^2 + 2a_1a_2 \cos 2\pi(\vartheta_1 - \vartheta_2)uv + a_2^2v^2]$$

is such that for a sufficiently small $\epsilon > 0$ and for some positive $\eta = \eta(\epsilon; \mathrm{P}^0)$ one has

(26) $$Q(u, v; \vartheta_1, \vartheta_2) \geqq (u^2 + v^2)\eta, \text{ if } \mid \vartheta_1 - \vartheta_1^0 \mid < \epsilon, \mid \vartheta_2 - \vartheta_2^0 \mid < \epsilon.$$

This follows for reasons of continuity from the fact that $Q(u, v; \vartheta_1^0, \vartheta_2^0)$ is positive definite in view of (24) and (25). On the other hand, if a point

(27) $$\mathrm{P}^* = (\vartheta_1^*, \cdots, \vartheta_n^*)$$

is on N, then it is seen from (13) and from (25) that

(28) $$[F(\vartheta_1^* + u, \vartheta_2^* + v, \vartheta_3^*, \cdots, \vartheta_n^*)]^2 + [G(\vartheta_1^* + u, \vartheta_2^* + v, \vartheta_3^*, \cdots, \vartheta_n^*)]^2 = Q(u, v; \vartheta_1^*, \vartheta_2^*) + o(u^2 + v^2) \text{ as } u^2 + v^2 \to 0,$$

and that the o-term holds uniformly for all choices of the point (27) on N. Hence it is seen from (26) that for a sufficiently small $\delta > 0$ and for some positive $\zeta = \zeta(\delta; \mathrm{P}^0)$ one has

(29) $$[F(\vartheta_1^* + u, \vartheta_2^* + v, \vartheta_3^*, \cdots, \vartheta_n^*)]^2 + [G(\vartheta_1^* + u, \vartheta_2^* + v, \vartheta_3^*, \cdots, \vartheta_n^*)]^2 \geqq (u^2 + v^2)\zeta$$

whenever (27) is a point of N such that

(30) $$\mid \vartheta_1^* - \vartheta_1^0 \mid < \delta, \mid \vartheta_2^* - \vartheta_2^0 \mid < \delta, \text{ while } u^2 + v^2 < \delta^2.$$

Now it is clear from (11) and (13) that

(31) $$\mid \chi(\vartheta_1, \cdots, \vartheta_n) \mid \leqq \text{const.} \{[F(\vartheta_1, \cdots, \vartheta_n)]^2 + [G(\vartheta_1, \cdots, \vartheta_n)]^2\}^{-\frac{1}{2}}$$

for all points $(\vartheta_1, \cdots, \vartheta_n)$ of Θ which do not lie on N. It is seen from (29) and (31) that

(32) $$\mid \chi(\vartheta_1^* + u, \vartheta_2^* + v, \vartheta_3^*, \cdots, \vartheta_n^*) \mid = O([u^2 + v^2]^{-\frac{1}{2}}), \ u^2 + v^2 \to 0,$$

holds uniformly with respect to P^*, if P^0 is fixed and (30) is satisfied. This clearly implies that the contribution of a sufficiently small vicinity of P^0 to the n-fold integral (22) is finite. Since P^0 is any point of the closed, bounded,

$(n-2)$-dimensional manifold (13), and since this manifold consists of the zeros of the denominator of (11), the proof of (22) is complete.

The admissibility of the identification of the space-average with the time-average as mentioned at the beginning of the paper may be treated as follows: Consider the transformation

$$\tau_t = \tau_t(\vartheta_1, \cdots, \vartheta_n)$$

which sends a point

$$(\vartheta_1, \cdots, \vartheta_n)$$

of Θ into the point

$$(\lambda_1 t + \vartheta_1, \cdots, \lambda_n t + \vartheta_n)$$

of Θ. Thus τ_t is a measure-preserving transformation of Θ into itself, satisfies the group condition $\tau_r \tau_s = \tau_{r+s}$ and is of the metrically transitive type [8] in view of the linear independence of the λ_k. Hence the ergodic theorem of Birkhoff [7] is applicable to every L-integrable function $\nu = \nu(\vartheta_1, \cdots, \vartheta_n)$ on Θ and thus, according to (22), to the function $\nu = \chi$. Hence on excluding from Θ a set of points $(\vartheta_1, \cdots, \vartheta_n)$ of measure zero, the time-average (3) of the function

$$g(t) = \chi(\lambda_1 t + \vartheta_1, \cdots, \lambda_n t + \vartheta_n)$$

exists and is equal to the integral of the function (11) over the torus Θ. This means in view of (15), (17) and (1) that on keeping the frequencies λ_k and the amplitudes a_k fixed and on excluding from the n-dimensional space of the phases α_k a set of measure zero, (7) holds for the derivative $\phi'(t)$ of the angular function $\phi(t)$ of (1) and for the asymptotic distribution function $\sigma(\xi)$. Finally, (3) follows from (7) in view of (8).

It may be mentioned that on excluding, for fixed a_k and λ_k, a set of phases $(\alpha_1, \cdots, \alpha_n)$ which is $(n-1)$-dimensional, hence of measure zero, the function (1) is distinct from zero for every t, so that one can choose in (2) the sign $+$ for every t, in which case $|z(t)|$ and $\phi(t)$ become polar coördinates in the (x, y)-plane. It is known [9] that, in virtue of the linear independence of the λ_k, the asymptotic distribution function of the polar *angle* $\phi(t)$ thus defined is a circular equidistribution, since the asymptotic distribution of (1) in the (x, y)-plane is of radial symmetry.

THE JOHNS HOPKINS UNIVERSITY.

BIBLIOGRAPHY.

[1] G. D. Birkhoff, "Proof of the ergodic theorem," *Proceedings of the National Academy of Sciences*, vol. 17 (1931), pp. 656-660.

[9] Wintner [10].

[2] G. D. Birkhoff, "Probability and physical systems," *Bulletin of the American Mathematical Society*, vol. 38 (1932), pp. 361-379.

[3] P. Bohl, "Über ein in der Theorie der säkularen Störungen vorkommendes Problem," *Journal für die reine und angewandte Mathematik*, vol. 135 (1909), pp. 189-283.

[4] G. H. Hardy and J. E. Littlewood, "On Lindelöf's hypothesis concerning the Riemann zeta-function," *Proceedings of the Royal Society*, ser. A, vol. 103 (1923), pp. 403-412.

[5] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 48-88.

[6] A. Khintchine, "Zu Birkhoffs Lösung des Ergodenproblems," *Mathematische Annalen*, vol. 107 (1933), pp. 485-488.

[7] H. Weyl, "Sur une application de la théorie des nombres à la mécanique statistique et la théorie des perturbations," *Enseignement Mathématique*, vol. 16 (1914), pp. 455-467.

[8] A. Wintner, "Diophantische Approximationen und Hermitesche Matrizen," *Mathematische Zeitschrift*, vol. 30 (1929), pp. 290-319, more particularly pp. 312-316.

[9] A. Wintner, "Sur l'analyse anharmonique des inégalites séculaires fournies par l'approximation de Lagrange," *Rendiconti della R. Accademia Nazionale dei Lincei*, ser. 6, vol. 11 (1930), pp. 464-467.

[10] A. Wintner, "Upon a statistical method in the theory of diophantine approximations," *American Journal of Mathematics*, vol. 55 (1933), pp. 309-331.

ERRATA.

In the paper of E. R. van Kampen and Aurel Wintner, "On a symmetrical canonical reduction of the problem of three bodies," *American Journal of Mathematics*, vol. 59 (1937), pp. 153-166, read 18 instead of 9 in formula (55_2) on p. 165 and in formula (59) on p. 166.

In the paper of E. R. van Kampen and Aurel Wintner, "Convolutions of distributions on convex curves and the Riemann zeta function," *American Journal of Mathematics*, vol. 59 (1937), pp. 175-204,

page	line	instead of	read
188	14	$\epsilon = \bar{\epsilon}(\epsilon)$	$\bar{\epsilon} = \bar{\epsilon}(\epsilon)$
192	15	tangent	normal
192	16	the	the normal
192	18	$p_m{}^{-2\sigma})^2/$	$p_m{}^{-\sigma})^2/$
193	1	h	k
193	14	$<$	$>$
194	12	Max(	Max($\bar{\sigma}$,

ON AN ABSOLUTE CONSTANT IN THE THEORY OF VARIATIONAL STABILITY.[1]

By E. R. van Kampen and Aurel Wintner.

If $p(t)$, $-\infty < t < +\infty$, is a real continuous periodic function, the linear differential equation

$$\frac{d^2x}{dt^2} + p(t)x = 0 \tag{1}$$

is known to possess two solutions $x = x_1$, $x = x_2$ of the form

$$x_1 = e^{\lambda t/T} f_1(t), \qquad x_2 = e^{-\lambda t/T} f_2(t), \tag{2}$$

where $f_1(t)$ and $f_2(t)$ do not vanish identically and are periodic with the same period T as $p(t)$. The numbers λ and $-\lambda$, the characteristic exponents, are determined mod $2\pi i$ by the reciprocal quadratic equation

$$e^{\pm 2\lambda} - 2Ae^{\pm\lambda} + 1 = 0, \tag{3}$$

this equation being the characteristic equation of a real binary linear substitution of determinant 1. If λ is not a multiple of $T\pi i$, it is clear that the two solutions (2) are linearly independent. If λ is a multiple of $T\pi i$, it depends on the elementary divisors of that binary substitution whether or not the general solution of (1) is free of secular terms. The equation (1) determined by the given periodic function $p(t)$ is said to be of the stable type if every solution $x(t)$ remains bounded as $t \rightarrow \pm\infty$, i. e., if the elementary divisors are simple and λ lies on the imaginary axis of the λ-plane. Since from (3)

$$e^{\pm\lambda} = A \pm (A^2 - 1)^{\frac{1}{2}},$$

it follows that

$$-1 \leqq A \leqq 1 \tag{4}$$

is necessary and that

$$-1 < A < 1 \tag{5}$$

is sufficient for the stability of (1). In fact, a multiple elementary divisor is impossible unless there is a double root, so that $A^2 = 1$ is a necessary condition for solutions with secular terms.

The determination of the solutions (2) in terms of $p(t)$ requires, in

[1] Received February 22, 1937.

general, an application of infinite determinants or of equivalent transcendental processes. Even the determination of the characteristic exponents $\pm\lambda$ is quite involved, it being defined by the zeros of Hill's fundamental determinant or by the characteristic equation (3) of the monodromy matrix, i. e., by the number A. The latter can be represented, according to Liapounoff,[2] by means of the convergent series

$$A = 1 - A_1 + A_2 - A_3 + \cdots, \tag{6}$$

where, if $T > 0$ denotes the period of $p(t)$, the number A_n is the definite integral

$$A_n = \tfrac{1}{2}\int_0^T dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} Q_n(t_1,\cdots,t_n)\,dt_n, \tag{7}$$

while the function Q_n is defined in terms of the primitive function

$$P(t) = \int_0^t p(\bar{t})\,d\bar{t} \tag{8}$$

of the coefficient $p(t)$ of (1) as follows:

$$\begin{aligned} (9)\quad & Q_n(t_1,\cdots,t_n) \\ &= \{P(T) - P(t_1) + P(t_n)\}\{P(t_1) - P(t_2)\}\cdots\{P(t_{n-1}) - P(t_n)\}. \end{aligned}$$

In particular, Q_1 is the constant $P(T)$, so that, from (7) and (8),

$$A_1 = \tfrac{1}{2}TP(T) = \tfrac{1}{2}T\int_0^T p(t)\,dt. \tag{10}$$

Since the actual determination of $\pm\lambda$ requires, in view of (3), (6), (7), (8), (9), highly complicated operations, it is natural to ask for criteria which impose a less remote condition on the coefficient $p(t)$ of (1) and assure that (1) is of the stable type. First, if $p(t)$ is a positive constant, (1) is clearly of the stable type. On the other hand, if the positive periodic function $p(t)$ is not a constant, (1) need not be of the stable type. This holds also when the deviation of $p(t) > 0$ from a constant is less than an arbitrarily small $\epsilon > 0$. Examples to this effect are implied[3] by the theory of Mathieu's equation, where

$$p(t) = c_1 + c_2 \cos t \qquad (c_1 > c_2 \to 0 < c_1). \tag{11}$$

---

[2] A. Liapounoff, "Sur une série relative à la théorie des équations différentielles linéaires à coefficients périodiques," *Comptes Rendus*, vol. 123 (1896), pp. 1248-1252. This investigation of Liapounoff is reproduced on pp. 425-431 of vol. IV, part III (1902) of Forsyth's *Theory of Differential Equations*.

[3] Cf. M. J. O. Strutt, "Wirbelströme im elliptischen Zylinder," *Annalen der Physik*, vol. 84 (1927), pp. 485-506, where further references are given.

Now there exists, according to Liapounoff,[2] an absolute constant $\alpha > 0$ such that if a real continuous function $p(t)$ of period T is non-negative for every t, positive for some t and satisfies the inequality

$$(12) \qquad T\int_0^T p(t)\,dt \leqq \alpha,$$

then (1) is necessarily of the stable type. That α cannot be arbitrarily large, is clear from the example of the Mathieu equation mentioned above. It is indicated by Strutt's diagram,[3] which has been calculated numerically for the boundary curves of the stability region of (11) in the (c_1, c_2)-plane, that the best possible value of the absolute constant α cannot be much greater than 5. On the other hand, every $\alpha \leqq 4$ is admissible. In order to see this, it is, according to Liapounoff, sufficient to observe that, as shown by (8), (9), (7) and (10), the assumption

$$(13) \qquad 0 \leqq p(t) \not\equiv 0$$

obviously implies the inequalities

$$(14) \qquad A_{n+1} < \frac{A_n A_1}{n+1}, \qquad A_n > 0.$$

In fact, if (12) is satisfied by an $\alpha \leqq 4$, then

$$2 \geqq A_1 > A_2 > A_3 > \cdots > 0$$

in view of (10) and (14), so that the number (6) satisfies the sufficient condition (5) of stability.

Now it will be shown that $\alpha = 4$ is the exact value of the absolute constant in question. Thus there does or does not exist a continuous periodic function $p(t)$ which satisfies (13), (12) and makes (1) a differential equation of the unstable type according as $\alpha > 4$ or $\alpha \leqq 4$. It also will be shown that $\alpha = 4$ remains the greatest admissible value of α also when one restricts $p(t)$ to be an even function of t. Finally, it will be seen from the proof that nothing is gained if one requires that $p(t)$ is analytic or if one replaces (13) by $p(t) > 0$.

Needless to say, the greatest admissible value of α is independent of the value of the period T. This is seen from (12) if one replaces t in (1) by ct, where $c > 0$ is arbitrary. On choosing

$$(15) \qquad T = 1,$$

it follows that, since every $\alpha \leqq 4$ is admissible, the statement to be proved may be formulated as follows: There exists for every $\epsilon > 0$ a real non-negative

continuous function $p(t) \not\equiv 0$ of period 1 such that, on the one hand, the mean value of $p(t)$ satisfies the inequality

$$\int_0^1 p(t)\,dt \leqq 4(1+\epsilon) \tag{16}$$

and, on the other hand, the characteristic exponents $\pm\lambda$ of (1) are not of the stable type.

For two fixed numbers β, μ which satisfy the inequalities

$$0 < \mu < \tfrac{1}{2}, \qquad \beta > 0, \tag{17}$$

put

$$\begin{aligned} &p(t) = (\mu - t)\beta/\mu, \text{ if } 0 \leqq t \leqq \mu \text{ and } p(t) = 0, \text{ if } \mu \leqq t \leqq \tfrac{1}{2},\\ &p(t) = p(1-t), \text{ if } \tfrac{1}{2} < t < 1, \text{ finally } p(t+1) = p(t), \end{aligned} \tag{18}$$

so that $p(t) \not\equiv 0$ is an even, continuous, non-negative function of period 1. It is easy to see that, for this $p(t)$,

$$A_1 = \tfrac{1}{2}\beta\mu \tag{19}$$

and

$$A_2 < \beta^2\mu^3. \tag{20}$$

In fact, (19) is clear from (10), (15) and (18). Furthermore, from (18),

$$p(t) = 0, \text{ if } \mu \leqq t \leqq 1 - \mu;$$

hence, from (8),

$$P(t_1) - P(t_2) = 0, \text{ if } \mu \leqq t_2 \leqq t_1 \leqq 1 - \mu,$$

and so, since

$$Q_2(t_1, t_2) = \{P(1) - P(t_1) + P(t_2)\}\{P(t_1) - P(t_2)\} \tag{21}$$

in view of (9) and (15),

$$Q_2(t_1, t_2) = 0, \text{ if } \mu \leqq t_2 \leqq t_1 \leqq 1 - \mu. \tag{22}$$

Since (8) is a non-decreasing function in view of $p(t) \geqq 0$, it is seen from (21) that

$$Q_2(t_1, t_2) \leqq \{P(1)\}^2, \quad \text{whenever} \quad 0 \leqq t_2 \leqq t_1 \leqq 1.$$

Since $P(1) = \beta\mu$ in view of (10), (15) and (19), it follows that

$$Q_2(t_1, t_2) \leqq \beta^2\mu^2, \quad \text{whenever} \quad 0 \leqq t_2 \leqq t_1 \leqq 1. \tag{23}$$

Now if S denotes that portion of the triangle $0 \leqq t_2 \leqq t_1 \leqq 1$ in the (t_1, t_2)-

plane on which $\mu \leqq t_2 \leqq t_1 \leqq 1 - \mu$ does not hold, then it is seen from (7), (15), (22) and (23) that

$$2A_2 = \int_0^1 \int_0^{t_1} Q_2(t_1, t_2)\, dt_2 dt_1 = \int\int_S Q_2(t_1, t_2)\, dt_2 dt_1 \leqq \beta^2\mu^2 \int\int_S dt_2 dt_1.$$

This proves (20), since, by the definition of the region S,

$$\int\int_S dt_1 dt_2 = \text{area of } S < 2\mu.$$

Now let $\epsilon > 0$ in (16) be given. Since only small values of ϵ need to be considered, one can assume that $\epsilon < \frac{1}{2}$ and then choose the numbers β and μ, which occur in (17) and define the periodic peak function $p(t) \geqq 0$, in such a way that on the one hand

$$\beta\mu = 4(1 + \epsilon) \tag{24}$$

and on the other hand $\mu < \epsilon/(4 + 4\epsilon)^2$. Since the latter inequality implies, in view of (20), that

$$A_2 < \beta^2\mu^2\epsilon/(4 + 4\epsilon)^2,$$

it is seen from (24) and from the assumption $\epsilon < \frac{1}{2}$ that

$$A_2 < \epsilon < \tfrac{1}{2}. \tag{25}$$

On the other hand,

$$A_1 = 2(1 + \epsilon) \tag{26}$$

in view of (19) and (24). Since, by (25), (26) and (14),

$$A_2 > A_3 > A_4 > \cdots > 0,$$

it is clear that

$$0 < A_2 - A_3 + A_4 - \cdots < A_2.$$

Hence $A < 1 - A_1 + A_2$ in view of (6). It follows, therefore, from (25) and (26) that

$$A < 1 - A_1 + \epsilon = 1 - 2(1 + \epsilon) + \epsilon = -1 - \epsilon < -1.$$

Consequently, (4) is not satisfied. Since (4) is a necessary condition for characteristic exponents $\pm \lambda$ of the stable type, and since (16) is satisfied in view of (10), (15) and (26), the proof is complete.

THE JOHNS HOPKINS UNIVERSITY.

ON THE EXPANSION OF THE REMAINDER IN THE OPEN-TYPE NEWTON-COTES QUADRATURE FORMULA.*

By Orville G. Harrold, Jr.

1. The result that the remainder in the Newton-Cotes quadrature formula can be expanded in a series of the Euler-MacLaurin type has been established by J. V. Uspensky.[1] Inasmuch as the open-type quadrature formula of this kind discussed by J. F. Steffensen [2] is important in the numerical integration of differential equations, the question has been raised as to whether or not an analogous development holds for the open formula. The answer is in the affirmative.

2. We consider, without loss of generality, the integration interval $(0, 1)$. The unit interval is divided in n equal parts. The function $f(x)$ to be integrated over this interval is assumed known at $x = 1/n, 2/n, \cdots, (n-1)/n$. The coefficients, or weights, in the quadrature formula will be denoted by A_i:

$$A_i = \int_0^1 \frac{\omega_n(x)\,dx}{\omega'_n(i/n)(x - i/n)},$$
$$\omega_n(x) = (x - 1/n)(x - 2/n) \cdots (x - (n-1)/n).$$

It is convenient to introduce the symbol

$$K_n{}^\nu = A_1 B_\nu(1/n) + A_2 B_\nu(2/n) + \cdots + A_{n-1} B_\nu((n-1)/n),$$

where $B_\nu(x)$ is the Bernoullian polynomial of degree ν.

3. With the above notations, the result of this investigation can be formulated more explicitly as follows:

Let $f(x)$ be a continuous function on $0 \leqq x \leqq 1$, with as many continuous derivatives as are needed in the discussion; *then*

* Received November 10, 1936; revised December 28, 1936.

[1] J. V. Uspensky, " On the expansion of the remainder in the Newton-Cotes formula," *Transactions of the American Mathematical Society*, vol. 37 (1935), pp. 381-396.

[2] J. F. Steffensen, *Interpolation.* Williams and Wilkins Co., Baltimore, 1927.

$$\int_0^1 f(x)\,dx = \sum_{i=1}^{n-1} A_i f(i/n) - \sum_{\nu=m=[n/2]}^{m+s-1} K_n^{2\nu} \frac{\Delta f^{(2\nu-1)}(0)}{(2\nu)!} + R_{2s+2m},$$

where

$$R_{2s+2m} = -K_n^{2s+2m} \frac{f^{(2s+2m)}(\xi)}{(2s+2m)!}, \qquad 0 < \xi < 1,$$

s being an arbitrary positive integer. Further, if the even ordered derivatives keep their sign on $0 \leqq x \leqq 1$ and if their signs are all alike, then when the series is truncated after a certain term, the error committed is of the same sign as the next term and in absolute value is less than it.

4. By the Euler-Maclaurin summation formula

$$f(x+\theta) = \int_x^{x+1} f(t)\,dt + \sum_{\nu=1}^{r} \frac{B_\nu(\theta)}{\nu!} \Delta f^{(\nu-1)}(x) - \int_0^1 \frac{\bar{B}_r(\theta - t)}{r!} f^{(r)}(x+t)\,dt, \qquad (0 \leqq \theta \leqq 1).$$

$\bar{B}_\nu(x)$ is the Bernoullian periodic function of order ν, r being an arbitrary positive integer. Fixing $x = 0$ throughout this discussion and allowing θ to take on successively the values $1/n, 2/n, \cdots, (n-1)/n$, we get

$$f(i/n) = \int_0^1 f(t)\,dt + \sum_{\nu=1}^{r} \frac{B_\nu(i/n)}{\nu!} \Delta f^{(\nu-1)}(0) - \int_0^1 \frac{\bar{B}_r(i/n - t)}{r!} f^{(r)}(t)\,dt, \qquad (i = 1, 2, \cdots, n-1).$$

Multiplying by A_i and summing from $i = 1$ to $n-1$, there results (since $\sum_{i=1}^{n-1} A_i = 1$),

$$\int_0^1 f(t)\,dt = \sum_{i=1}^{n-1} A_i f(i/n) - \sum_{\nu=1}^{r} \frac{K_n^\nu}{\nu!} \Delta f^{(\nu-1)}(0) + \int_0^1 \frac{f^{(r)}(t)}{r!} \{A_1 \bar{B}_r(1/n - t) + A_2 \bar{B}_r(2/n - t) + \cdots + A_{n-1} \bar{B}_r((n-1)/n - t)\}\,dt.$$

For $\nu \leqq 2m - 1$, $m = [n/2]$

$$(*) \qquad K_n^\nu = \int_0^1 B_\nu(x)\,dx = 0.$$

Since $A_i = A_{n-i}$ and $B_{2\nu+1}(i/n) = -B_{2\nu+1}((n-i)/n)$ for all values of ν, we have

$$K_n^{2\nu+1} = 0.$$

Thus the integral may be presented, for $r = 2s$,

$$\int_0^1 f(t)\,dt = \sum_{k=1}^{n-1} A_k f(k/n) - \sum_{\nu=m}^{s-1} K_n{}^{2\nu} \frac{\Delta f^{(2\nu-1)}(0)}{(2\nu)!} + R_{2s},$$

where

$$R_{2s} = \int_0^1 \frac{f^{(2s)}(t)}{(2s)!} \{ \sum_{i=1}^{n-1} A_i (\bar{B}_{2s}(i/n - t) - B_{2s}(i/n) \} dt.$$

To establish the result asserted in § 3 it suffices to show:

1°, the numbers $K_n{}^{2m}, K_n{}^{2m+2}, \cdots$ alternate in sign;

2°, the quantity $\sum_{i=1}^{n-1} A_i \{ \bar{B}_{2s}(i/n - t) - B_{2s}(i/n) \}$ keeps its sign on $0 < t < 1$.

For if 2° is true, we may write, by virtue of (*)

$$R_{2s} = -\frac{f^{(2s)}(\xi)}{(2s)!} \cdot K_n{}^{2s}, \qquad 0 < \xi < 1.$$

5. To establish the points in question the following properties of

$$Q_k(t) \equiv \sum_{i=1}^{n-1} A_i \{ \bar{B}_k(i/n - t) - B_k(i/n) \}$$

are noted:

(1) $Q_k(t) = (-1)^k Q_k(1 - t)$;

(2) $Q_k(t)$ is continuous for $k = 2, 3, \cdots$, $Q_k(t)$ possesses derivatives of orders $1, 2, \cdots, k - 1$, $Q_k{}^{(k-2)}(t)$ is not differentiable at $t = i/n$;

(3) $$Q''_{2k-1}(t) = (2k-1)(2k-2) Q_{2k-3}(t), \qquad (k = 2, 3, \cdots (t \neq i/n,\ k = 2)),$$
$$Q'_{2k}(t) = -2k Q_{2k-1}(t) \qquad (k = 1, 2, \cdots (t \neq i/n,\ k = 1)),$$
$$Q''_{2k}(t) = 2k(2k-1) \{ Q_{2k-2}(t) + \sum_{i=1}^{n-1} A_i B_{2k-2}(i/n) \} \qquad (k = 2, 3, \cdots);$$

(4) $Q_k(0) = Q_k(1) = 0, \qquad (k = 1, 2, \cdots)$.

Let α_k and β_k denote respectively the number of distinct zeros of $Q_{2k-1}(t)$ and $Q_{2k}(t)$ on $0 < t < 1$. By virtue of (1), for $k = 1, 2, \cdots$,

$$\alpha_k \geqq 1. \tag{5}$$

From the fact that $Q_{2k}(t)$ has $\beta_k + 2$ distinct zeros on $0 \leqq t \leqq 1$, we get, by the use of (3) and Rolle's theorem, that $Q_{2k-1}(t)$ has at least $\beta_k + 1$ distinct zeros on $0 < t < 1$; thus,

$$\beta_k + 1 \leqq \alpha_k. \tag{6}$$

Due to the fact that $Q_{2k-1}(t)$ has $\alpha_k + 2$ distinct zeros on $0 \leqq t \leqq 1$, we get by (3) and repeated application of Rolle's theorem that $Q_{2k-3}(t)$ has at least α_k distinct zeros on $0 < t < 1$; hence,

$$\alpha_k \leqq \alpha_{k-1}. \tag{7}$$

If it can be shown that $\alpha_m = 1$, then, by (6) and (7) $\beta_m, \beta_{m+1}, \cdots$ are all zero. Thus, one of our contentions, namely, that

$$\sum_{i=1}^{n-1} A_i\{\bar{B}_{2s}(i/n - t) - B_{2s}(i/n)\}$$

keeps its sign on $0 < t < 1$, will be established.

6. Consider the function $Q_{2m-1}(t) \equiv \sum_{i=1}^{n-1} A_i\{\bar{B}_{2m-1}(i/n - t) - B_{2m-1}(i/n)\}$. Since $\sum_{i=1}^{n-1} A_i B_{2m-1}(i/n) = \int_0^1 B_{2m-1}(x)\,dx = 0$ and $B_m(u+1) - B_m(u) = mu^{m-1}$ we may make the following simplifications:

$$Q_{2m-1}(t) = \sum_{i=1}^{n-1} A_i \bar{B}_{2m-1}(i/n - t) = \sum_{i=1}^{i/n \leqq t} A_i \bar{B}_{2m-1}(i/n - t) + \sum_{i/n > t}^{n-1} A_i \bar{B}_{2m-1}(i/n - t);$$

now

$$\begin{aligned}\sum_{i/n > t}^{n-1} A_i B_{2m-1}(i/n - t) &= \sum_{i/n > t}^{n-1} A_i B_{2m-1}(i/n - t)\\ &= \sum_{i=1}^{n-1} A_i B_{2m-1}(i/n - t) - \sum_{i/n \leqq t} A_i B_{2m-1}(i/n - t),\end{aligned}$$

hence

$$\begin{aligned}Q_{2m-1}(t) &= \sum_{i=1}^{n-1} A_i B_{2m-1}(i/n - t) + \sum_{i/n \leqq t} A_i\{B_{2m-1}(1 + i/n - t) - B_{2m-1}(i/n - t)\}\\ &= \sum_{i=1}^{n-1} A_i B_{2m-1}(i/n - t) + \sum_{i/n \leqq t} A_i (i/n - t)^{2m-2}(2m-1).\end{aligned}$$

The first term on the right is $\int_0^1 B_{2m-1}(x - t)\,dx = -t^{2m-1}$. By comparison of

$$Q_{2m-1}(t) = -t^{2m-1} + (2m-1)\sum_{i/n \leqq t} A_i (i/n - t)^{2m-2}$$

with

$$\int_0^1 \phi(t)\,dt = \sum_{i=1}^{n-1} A_i\phi(i/n) + R$$

we see that $R_0(t) \equiv -Q_{2m-1}(t)/(2m-1)$ is precisely the remainder obtained when the open formula is applied to

$$\phi(x) = \begin{cases} (x-t)^{2m-2}, & x \leqq t, \quad 0 < t < 1, \\ 0, & x > t. \end{cases}$$

Setting

$$R_k(t) = \frac{(-1)^k t^{2m-k-1}}{2m-k-1} - \sum_{i/n \leqq t} A_i(i/n - t)^{2m-k-2},$$

let the number of distinct zeros of $R_k(t)$, for $k = 0, 1, \cdots, 2m-3$ be denoted by $N_0, N_1, \cdots, N_{2m-3}$. Let the number of variations in sign of $R_{2m-2}(t)$, as t varies from 0 to 1 be denoted by N_{2m-2}. Evidently $R_k(t)$ is continuous $(k = 0, 1, 2m-3)$ and possesses a derivative $R'_k(t) = -(2m-k-2)R_{k+1}(t)$, $(k = 0, 1, \cdots, 2m-4)$; but $R_{2m-3}(t)$ does not possess a derivative at $t = i/n$.

By the property of the open quadrature formula

$$R_k(0) = R_k(1) = 0;$$

hence, $R_k(t)$ has $N_k + 2$ distinct zeros on $0 \leqq t \leqq 1$, so that, by Rolle's theorem, and the fact that $R'_k(t) = -(2m-k-2)R_{k+1}(t)$ we get

$$\text{(A)} \qquad N_k + 1 \leqq N_{k+1}$$

from which it follows that

$$N_0 + (2m-3) \leqq N_{2m-3}.$$

In particular, $N_{2m-3} + 1 \leqq N_{2m-2}$. From the fact that the coefficients A_i alternate in sign for $i \leqq m$,[3] it follows that

$$R_{2m-2}(t) = t - \Sigma A_i, \qquad i/n \leqq t,$$

can have at most $2m-1$ variations in sign in $0 < t < 1$ when $n = 2m$. For, we note $R_{2m}(t)$ cannot change sign more than twice in each of the subintervals $(2i-1)/n \leqq t < (2i+1)/n$, while it definitely does not change sign in $0 \leqq t < 1/n$, and has at most one change of sign in $(2m-1)/2m \leqq t < 1$. Hence, for even n,

$$N_{2m-2} \leqq 2m-1,$$

[3] See Lemma I below $(n \neq 5)$.

from which $N_{2m-3} \leqq 2m - 2$, or $N_0 \leqq +1$. But $N_0 = \alpha_m \geqq +1$; thus $\alpha_m = +1$, $\beta_m = \beta_{m+1} = \cdots = 0$.

If $n = 2m + 1$ we use the relations

$$R_{2m-2}(t) = t - \Sigma A_i, \qquad i/n \leqq t,$$
$$R_k(t) = (-1)^{k-1} R_k(1 - t),$$

which follow immediately from the definition of $R_k(t)$.

Two cases are distinguished:

1°. $n = 4k + 1$. As before, $R_{2m-2}(t)$ does not change sign in $0 \leqq t < 1/n$, and not more than twice in each $(2i - 1)/n \leqq t < (2i + 1)/n$, $i = 1, 2, \cdots, k - 1$. It changes sign at most once in $(m - 1)/n \leqq t < m/n$ since $A_m = A_{2k}$ is negative by Lemma 1 and $R_{2m-2}(m/n)$ is obviously negative. Thus $R_{2m-2}(t)$ has at most $4(k - 1) + 2 + 1 = 2m - 1$ alternations of sign on $0 < t < 1$; hence, $N_{2m-2} \leqq 2m - 1$, $N_0 \leqq +1$, and as before, $N_0 = +1$.

2°. $n = 4k + 3$. Again, $R_{2m-2}(t)$ does not change sign in $0 \leqq t < 1/n$, and not more than twice in each $(2i - 1)/n \leqq t < (2i + 1)/n$. However, three alternations are possible in $(m - 1)/n \leqq t \leqq (m + 1)/n$ so that $N_{2m-2} \leqq 2m + 1$. It follows that $R_{2m-3}(t)$ has at most $2m$ zeros on $0 < t < 1$. But $R_{2m-3}(t)$ is an even function of t with respect to $t = 1/2$, so if $R_{2m-3}(t) < 0$ at $t = 1/2$, $N_{2m-3} \equiv 0 \pmod 4$; but $N_{2m-3} \leqq 2m = 4k + 2$, so that $N_{2m-3} \leqq 2m - 2$, hence $N_0 = +1$.

If $R_{2m-3}(1/2) \nless 0$, an impossible situation arises. Since

$$-R_{2m-3}(t) = t^2/2 + \sum_{i/n \leqq t} A_i(i/n - t),$$

our assumption implies that $\Sigma A_i(i/n) \leqq 1/8$, $i/n < 1/2$, which is $(n \neq 3)$ false by Lemma 2 below. Thus in all cases $\alpha_m = 1$, $\beta_m = 0$ $(n \neq 3, 5)$.[4]

7. The functions $Q_{2m}(t), Q_{2m+2}(t), \cdots$ do not change sign on $0 \leqq t \leqq 1$. They are periodic functions with continuous derivatives, and such that $Q'_{2\lambda}(0) = 0$, so that $Q_{2\lambda}(t)$ has the same sign as $Q''_{2\lambda}(0)$. From § 5,

$$Q''_{2m+2k}(0) = \sum_{i=1}^{n-1} A_i B_{2m+2k-2}(i/n),$$

and

$$\int_0^1 Q_{2m+2k}(t)\,dt = -\sum_{i=1}^{n-1} A_i B_{2m+2k}(i/n).$$

[4] 9 deals with the expansions for $n = 3, 5$.

It is thus evident that the coefficients $K_n^{2\nu}$ alternate in sign for fixed n, and $\nu = m, m+1, \cdots$. Both assertions of § 4 have now been established.

8. It remains to establish the two lemmas mentioned in § 6.

LEMMA 1. *The coefficients A_i satisfy the following inequalities for $i \leqq n/2$:*

$$A_i > 0, \qquad i \text{ odd},$$
$$A_i < 0, \qquad i \text{ even},$$

while for $i > n/2$ we use the fact that $A_i = A_{n-i}$.

To demonstrate that the A_i alternate in sign it is convenient to modify the notation slightly to indicate the dependence on n. Set

$$A_k = A_{nk} = (1/n)\int_0^n \frac{P_k(x)}{P_k(k)}\,dx, \qquad (k = 1, 2, \cdots, n-1),$$

where

$$P_k(x) = \frac{(x-1)(x-2)\cdots(x-n+1)}{x-k}.$$

Since $P_k(k) = (-1)^{n-k-1}\Gamma(k)\Gamma(n-k)$,

$$A_{nk} = \frac{(-1)^{n-k-1}}{n\Gamma(k)\Gamma(n-k)}\int_0^n P_k(x)\,dx.$$

Setting $J_n = \int_0^n P_k(x)\,dx$, and recalling well-known formulas for Gamma functions, we find

$$J_n = \frac{(-1)^{n-1}\Gamma(n)}{\pi}\int_0^1 \xi^{-1}(1-\xi)^{n-1}d\xi \int_0^n e^{x\log\xi/(1-\xi)}\,\frac{\sin\pi x}{x-k}\,dx.$$

The last integral may be split up and presented as

$$\int_0^n e^{x\log\xi/(1-\xi)}\,\frac{\sin\pi x}{x-k}\,dx = (-1)^k(\xi/(1-\xi))^k\left\{\int_0^{n-k} e^{t\log\xi/(1-\xi)}\,\frac{\sin\pi t}{t}\,dt + \int_0^k e^{t\log(1-\xi)/\xi}\,\frac{\sin\pi t}{t}\,dt\right\}.$$

Using the formula

$$\int_0^h e^{\alpha x}\,\frac{\sin\pi x}{x}\,dx = \pi/2 + \arctan\alpha/\pi + \pi(-1)^{h-1}\int_{-\infty}^{\alpha}\frac{e^{hx}}{\pi^2+x^2}\,dx,$$

we get

$$\int_0^n e^{x\log\xi/(1-\xi)}\frac{\sin\pi x dx}{x-k}$$
$$= (-1)^k\pi(\xi/(1-\xi))^k\left\{1+(-1)^{n-k-1}\int_{-\infty}^{\log\xi/(1-\xi)}\frac{e^{(n-k)x}dx}{\pi^2+x^2}\right.$$
$$\left.+(-1)^{k-1}\int_{-\infty}^{-\log\xi/(1-\xi)}\frac{e^{kx}dx}{\pi^2+x^2}\right\}.$$

Carrying out the substitution in J_n, we get

$$J_n=(-1)^{n+k-1}\Gamma(k)\Gamma(n-k)$$
$$+\Gamma(n)\left\{\int_0^1\xi^{k-1}(1-\xi)^{n-k-1}d\xi\int_{-\infty}^{\log\xi/(1-\xi)}\frac{e^{(n-k)x}dx}{\pi^2+x^2}\right.$$
$$\left.+(-1)^n\int_0^1\xi^{k-1}(1-\xi)^{n-k-1}d\xi\int_{-\infty}^{-\log\xi/(1-\xi)}\frac{e^{kx}dx}{\pi^2+x^2}\right\}.$$

Since $A_{nk}=\dfrac{(-1)^{n-k-1}}{n}\,\dfrac{J_n}{\Gamma(k)\Gamma(n-k)}$, we can now write

$$A_{nk}=\frac{1}{n}+\frac{n-1}{n}C_{n-2}^{k-1}(-1)^{k+1}\left\{(-1)^n\int_0^1\xi^{k-1}(1-\xi)^{n-k-1}d\xi\int_{-\infty}^{\log\xi/(1-\xi)}\frac{e^{(n-k)x}dx}{\pi^2+x^2}\right.$$
$$\left.+\int_0^1\xi^{k-1}(1-\xi)^{n-k-1}d\xi\int_{-\infty}^{-\log\xi/(1-\xi)}\frac{e^{kx}dx}{\pi^2+x^2}\right\}.$$

The quantity within the brackets becomes, after change of variable in the inner integrals,

$$\int_0^1\xi^{-1}(1-\xi)^{n-1}d\xi\int_0^1\frac{t^{k-1}+(-1)^nt^{n-k-1}}{\pi^2+\log^2(((1-\xi)/\xi)t)}dt.$$

If k is odd and $\leqq n/2$, we see at once that $A_{nk}>0$ for n even or odd. If k is even,

$$A_{nk}=\frac{1}{n}-\left(\frac{n-1}{n}\right)C_{n-2}^{k-1}\left\{\int_0^1\xi^{-1}(1-\xi)^{n-1}d\xi\int_0^1\frac{t^{k-1}+(-1)^nt^{n-k-1}}{\pi^2+\log^2(((1-\xi)/\xi)t)}dt\right\};$$

and, since the integrand is everywhere non-negative for $k\leqq n/2$,

$$-nA_{nk}>-1+(n-1)C_{n-2}^{k-1}\int_0^{1/2}\xi^{-1}(1-\xi)^{n-1}d\xi\int_{\xi/(1-\xi)}^1\frac{t^{k-1}+(-1)^nt^{n-k-1}}{\pi^2+\log^2(((1-\xi)/\xi)t)}dt$$

or

$$-nA_{nk}>-1+(n-1)C_{n-2}^{k-1}\int_0^{1/2}\frac{\xi^{-1}(1-\xi)^{n-1}d\xi}{\pi^2+\log^2((1-\xi)/\xi)}\int_{\xi/(1-\xi)}^1\{t^{k-1}+(-1)^nt^{n-k-1}\}dt.$$

On integration of the inner integral

$$-nA_{nk} > -1 + (n-1)C_{n-2}^{k-1}\left\{\frac{1}{k} + \frac{(-1)^n}{n-k}\right\}\int_0^{1/2} \frac{\xi^{-1}(1-\xi)^{n-1}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)}$$
$$-(n-1)C_{n-2}^{k-1}\left\{\frac{1}{k}\int_0^{1/2} \frac{\xi^{k-1}(1-\xi)^{n-k}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)}\right.$$
$$\left. + \frac{1}{n-k}\int_0^{1/2} \frac{\xi^{n-k-1}(1-\xi)^{k}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)}\right\}.$$

The term

$$(n-1)C_{n-2}^{k-1}\left\{\frac{1}{k}\int_0^{1/2} \frac{\xi^{k-1}(1-\xi)^{n-k}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)} + \frac{1}{n-k}\int_0^{1/2} \frac{\xi^{n-k-1}(1-\xi)^{k}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)}\right\}$$

is less than

$$\frac{(n-1)}{\pi^2}C_{n-2}^{k-1}\left\{\frac{1}{k}\int_0^{1/2} \xi^{k-1}(1-\xi)^{n-k-1}d\xi + \frac{1}{n-k}\int_0^{1/2} \xi^{n-k-1}(1-\xi)^{k-1}d\xi\right\}$$

which is less than

$$\frac{(n-1)}{\pi^2}C_{n-2}^{k-1}\frac{1}{k}\left\{\int_0^{1/2} \xi^{k-1}(1-\xi)^{n-k-1}d\xi + \int_0^{1/2} \xi^{n-k-1}(1-\xi)^{k-1}d\xi\right\}$$

for $k \leqq n/2$. Hence the term

$$(n-1)\, C_{n-2}^{k-1}\left\{\frac{1}{k}\int_0^{1/2} \frac{\xi^{k-1}(1-\xi)^{n-k}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)} + \frac{1}{n-k}\int_0^{1/2} \frac{\xi^{n-k-1}(1-\xi)^{k}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)}\right\}$$

is less than $1/k\pi^2$. Since

$$\int_0^{1/2} \frac{\xi^{-1}(1-\xi)^{n-1}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)} \geqq \left(\frac{n}{1+n}\right)^n \int_0^{1/(1+n)} \frac{\xi^{-1}(1-\xi)^{-1}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)},$$

we have

$$\int_0^{1/2} \frac{\xi^{-1}(1-\xi)^{n-1}d\xi}{\pi^2 + \log^2((1-\xi)/\xi)} \geqq \left(\frac{n}{1+n}\right)^n \frac{1}{\pi}\arctan\left(\frac{\pi}{\log n}\right);$$

hence

$$-nA_{nk} > (n-1)C_{n-2}^{k-1}\left\{\frac{1}{k} + \frac{(-1)^n}{n-k}\right\}\left(\frac{n}{1+n}\right)^n \frac{1}{\pi}\arctan\left(\frac{\pi}{\log n}\right)$$

$$\text{(B)} \qquad -1 - \frac{1}{k\pi^2}\cdots.$$

By calculation, $A_{12,2}$ and $A_{13,2}$ are both negative quantities. Considering n odd and even separately, we see readily that the right side of (B) is positive for $n \geqq 12$, and for even $k \leqq n/2$. Thus for all $k \leqq n/2$, $(n \neq 5)$,

$$A_{n,2l+1} > 0, \qquad A_{n,2l} < 0.$$

For n less than 12 we refer to tables.[2]

LEMMA 2. $\sum_{1/n<\frac{1}{2}} A_i(i/n) > 1/8,\ n \equiv 3 \pmod 4$ $(n > 3)$.

To establish this inequality we start from an expression for A_{nk} used in the proof of the previous lemma:

$$A_{nk} = \frac{1}{n} + \frac{n-1}{n} C^{k-1}_{n-2} \int_0^1 \xi^{-1}(1-\xi)^{n-1} d\xi \int_0^1 \frac{(-t)^{k-1} + (-t)^{n-k-1}}{\pi^2 + \log^2(((1-\xi)/\xi)t)}\, dt.$$

It follows that

$$\text{(A)} \qquad \sum_{k=1}^{m} A_{nk}k/n = \frac{m(m+1)}{2(2m+1)^2} + \frac{2m}{(2m+1)^2} \int_0^1 \xi^{-1}(1-\xi)^{2m} d\xi$$
$$\times \int_0^1 \sum_{k=1}^{m} \frac{kC^{k-1}_{2m-1}\{(-t)^{k-1} + (-t)^{2m-k}\}}{\pi^2 + \log^2(((1-\xi)/\xi)t)}\, dt.$$

It suffices to show the double integral (second term) exceeds $\dfrac{1}{8(2m+1)^2}$. To this end we start with the identity, valid for $p + q = 1$,

$$p^m + C_m{}^1 p^{m-1} q + \cdots + C_m{}^{\mu} p^{m-\mu} q^{\mu} = \frac{\int_0^p x^{m-\mu-1}(1-x)^{\mu} dx}{\int_0^1 x^{m-\mu-1}(1-x)^{\mu} dx}.$$

Setting $x = \dfrac{-y}{1-y}$ in the integrand of the numerator on the right side, we obtain for $m = n-2$, $\mu = \dfrac{n-3}{2}$,

$$p^{n-2} + C^1{}_{n-2} p^{n-3} q + \cdots + C^{(n-3)/2}_{n-2} p^{(n-1)/2} q^{(n-3)/2}$$
$$= (-1)^{(n-1)/2} \frac{\int_0^{-(p/q)} \dfrac{y^{(n-3)/2} dy}{(1-y)^{n-1}}}{\int_0^1 x^{(n-3)/2}(1-x)^{(n-3)/2} dx},$$

whence, putting $p = -tq$ (so that $q = (1-t)^{-1}$),

$$(-1)^{n-2} q^{n-2} \sum_{k=0}^{(n-3)/2} C^k{}_{n-2} t^{n-k-2} (-1)^k$$
$$= (-1)^{(n-1)/2} \int_0^t \frac{y^{(n-3)/2} dy}{(1-y)^{n-1}} \div \int_0^1 x^{(n-3)/2}(1-x)^{(n-3)/2} dx.$$

Further simplification gives

$$\sum_{k=0}^{m-1} C^k{}_{2m-1}(-1)^k t^{2m-k-1} = mC^m{}_{2m-1}(1-t)^{2m-1}\int_0^t \frac{y^{m-1}dy}{(1-y)^{2m}},$$

which upon differentiation leads to

$$-\sum_{k=0}^{m-1}(k+1)C^k{}_{2m-1}(-1)^k t^{2m-1-k} = -2m\sum_{k=0}^{m-1} C^k{}_{2m-1}(-1)^k t^{2m-1-k}$$
$$+ mC^m{}_{2m-1}\left\{\frac{t^m}{1-t} - t(2m-1)(1-t)^{2m-2}\int_0^t \frac{y^{m-1}dy}{(1-y)^{2m}}\right\}.$$

This relation may be written as

$$-\sum_{k=0}^{m-1}(k+1)C^k{}_{2m-1}(-1)^k t^{2m-k-1}$$
$$= mC^m{}_{2m-1}\left\{\frac{t^m}{1-t} + (t-2m)(1-t)^{2m-2}\int_0^t \frac{y^{m-1}dy}{(1-y)^{2m}}\right\}.$$

To get the other terms, we notice

$$\sum_{k=0}^{2m-1} C^k{}_{2m-1}(-1)^k t^{2m-1-k} = (t-1)^{2m-1},$$

so that

$$\sum_{k=0}^{m-1} C^k{}_{2m-1}(-1)^k t^{k+1} = t(1-t)^{2m-1} + mC^m{}_{2m-1}t(1-t)^{2m-1}\int_0^t \frac{y^{m-1}dy}{(1-y)^{2m}},$$

and, again by differentiation

$$\sum_{k=0}^{m-1}(k+1)(-1)^k C^k{}_{2m-1}t^k = (1-t)^{2m-2}(1-2mt)$$
$$+ mC^m{}_{2m-1}\left\{(1-t)^{2m-2}(1-2mt)\int_0^t \frac{y^{m-1}dy}{(1-y)^{2m}} + \frac{t^m}{1-t}\right\}.$$

Denoting the polynomial in the inner integral in the formula for A_{nk} by $F(t)$, we have

$$\text{(C)}\quad F(t) = \sum_{k=0}^{m-1}(k+1)(-1)^k C^k{}_{2m-1}\{t^k - t^{2m-1-k}\},$$
$$= (1-t)^{2m-2}(1-2mt)$$
$$+ mC^m{}_{2m-1}\left\{\frac{2t^m}{1-t} - (2m-1)(1+t)(1-t)^{2m-2}\int_0^t \frac{y^{m-1}dy}{(1-y)^{2m}}\right\}.$$

Setting

$$\psi(t) = \frac{2t^m}{1-t} - (2m-1)(1+t)(1-t)^{2m-2}\int_0^t \frac{y^{m-1}dy}{(1-y)^{2m}},$$

it is evident that $\psi(0)=0$ and that $\psi(t)>0$ for $0<t<1$, since

$$\left\{\frac{\psi(t)}{(1-t)^{2m-2}(1+t)}\right\}'=\frac{t^{m-1}}{(1+t)^2(1-t)^{2m-2}}>0.$$

Thus, by integration,

$$\begin{aligned}\phi(t)=\int_0^t F(x)\,dx&=\int_0^t(1-x)^{2m-2}(1-2mx)\,dx+mC^m{}_{2m-1}\int_0^t\psi(x)\,dx,\\&=t(1-t)^{2m-1}+mC^m{}_{2m-1}\int_0^t\psi(x)\,dx>0,\end{aligned}$$

for $0<t<1$. This expression we use in the integral in (A):

$$\begin{aligned}\int_0^1\frac{F(t)\,dt}{\pi^2+\log^2(((1-\xi)/\xi)t)}&=\frac{\phi(1)}{\pi^2+\log^2((1-\xi)/\xi)}\\&\quad+\int_0^1\frac{\phi(t)2\log(((1-\xi)/\xi)t)}{[\pi^2+\log^2(((1-\xi)/\xi)t)]^2}\frac{dt}{t},\\&=\frac{\phi(1)}{\pi^2+\log^2((1-\xi)/\xi)}+R(\xi).\end{aligned}$$

If $1/2<\xi<1$, $\log((1-\xi)/\xi)t<0$; if $0<\xi\leqq 1/2$, $\log((1-\xi)/\xi)t<0$ for $t<\xi/(1-\xi)$. We seek now upper bounds for the absolute values of the negative terms:

$$\int_0^1\frac{\phi(t)2\log(((1-\xi)/\xi)t)}{[\pi^2+\log^2(((1-\xi)/\xi)t)]^2}\frac{dt}{t},\qquad 1/2<\xi<1;$$

$$\int_0^{\xi/(1-\xi)}\frac{\phi(t)2\log(((1-\xi)/\xi)t)}{[\pi^2+\log^2(((1-\xi)/\xi)t)]^2}\frac{dt}{t},\qquad 0<\xi\leqq 1/2.$$

Observing that

$$\begin{aligned}\phi(t)=t(1-t)^{2m-1}+mC^m{}_{2m-1}\Big\{&t^{m-1}\left(\frac{2}{2m-1}-\frac{1}{2m}\right)+t^m\left(\frac{m-1}{2m^2}-\frac{1}{m}\right)\\&+(m-1)(1-t)^{2m-1}\left(\frac{1-t}{2m}-\frac{2}{2m-1}\right)\int_0^t\frac{y^{m-2}dy}{(1-y)^{2m-1}}\Big\},\end{aligned}$$

we have, since the last two terms of the bracket are negative,

$$\phi(t)<t(1-t)^{2m-1}+\frac{2m+1}{2(2m-1)}C^m{}_{2m-1}t^{m-1}<\frac{1}{2m}+\frac{2m+1}{2(2m-1)}C^m{}_{2m-1},$$

for $0<t<1$. By virtue of these inequalities, the first integral is less in absolute value than

$$\frac{\dfrac{1}{2m}+\dfrac{2m+1}{2(2m-1)}C^m{}_{2m-1}}{\pi^2+\log^2\left(\dfrac{1-\xi}{\xi}\right)},$$

and the second is less than

$$\frac{1}{\pi^2}\left\{\frac{\xi}{1-\xi}+\frac{2m+1}{2(2m-1)}C^m{}_{2m-1}\left(\frac{\xi}{1-\xi}\right)^{m-1}\right\}.$$

Considering

$$\int_0^1 \xi^{-1}(1-\xi)^{2m}\left\{\frac{\phi(1)}{\pi^2+\log^2((1-\xi)/\xi)}+R(\xi)\right\}d\xi,$$

we see that this is greater than

$$\phi(1)\int_0^1\frac{\xi^{-1}(1-\xi)^{2m}d\xi}{\pi^2+\log^2((1-\xi)/\xi)}-\left(\frac{1}{2m}+\frac{2m+1}{2(2m-1)}C^m{}_{2m-1}\right)\int_{1/2}^1\frac{\xi^{-1}(1-\xi)^{2m}d\xi}{\pi^2+\log^2((1-\xi)/\xi)}$$
$$-\frac{1}{\pi^2}\int_0^1\xi^{-1}(1-\xi)^{2m}\left\{\frac{\xi}{1-\xi}+\frac{2m+1}{2(2m-1)}C^m{}_{2m-1}\left(\frac{\xi}{1-\xi}\right)^{m-1}\right\}d\xi,$$

which exceeds

$$\phi(1)\int_0^{1/2}\frac{\xi^{-1}(1-\xi)^{2m}d\xi}{\pi^2+\log^2((1-\xi)/\xi)}$$
$$-\frac{2m+1}{2(2m-1)}C^m{}_{2m-1}\int_{1/2}^1\frac{\xi^{-1}(1-\xi)^{2m}d\xi}{\pi^2+\log^2((1-\xi)/\xi)}-\frac{1}{\pi^2}\int_0^1(1-\xi)^{2m-1}d\xi$$
$$-\frac{(2m+1)}{2\pi^2(2m-1)}C^m{}_{2m-1}\int_0^1\xi^{m-2}(1-\xi)^{m+1}d\xi-\frac{1}{\pi^2 m(2m+1)2^{2m+1}}.$$

The second, third, and fourth terms are respectively not greater than

$$\frac{C^m{}_{2m-1}}{\pi^2(2m-1)2^{2m+1}},\quad \frac{1}{2\pi^2 m},\quad\text{and}\quad \frac{(m+1)(2m+1)}{4\pi^2 m(m-1)(2m-1)}.$$

Since

$$\int_0^{1/2}\frac{\xi^{-1}(1-\xi)^{2m}d\xi}{\pi^2+\log^2((1-\xi)/\xi)}\geqq\left(\frac{2m+1}{2m+2}\right)^{2m+1}\frac{1}{\pi}\operatorname{arc\,tan}\frac{\pi}{\log(2m+1)},$$

we have

$$\int_0^1\xi^{-1}(1-\xi)^{2m}d\xi\int_0^1\frac{F(t)\,dt}{\pi^2+\log^2(((1-\xi)/\xi)t)}$$
$$\geqq\frac{C^m{}_{2m-1}\left(\dfrac{2m+1}{2m+2}\right)^{2m+1}}{2m(2m-1)}\;\frac{\operatorname{arc\,tan}\dfrac{\pi}{\log(2m+1)}}{\pi}$$
$$-\frac{C^m{}_{2m-1}}{\pi^2(2m-1)2^{2m+1}}-\frac{1}{\pi^2 m}-\frac{(m+1)(2m+1)}{4\pi^2 m(m-1)(2m-1)}.$$

The coefficient of the integral is $2m/(2m+1)^2$; hence, we are concerned with the truth of the inequality

$$\frac{2m}{(2m+1)^2}\left\{C^m{}_{2m-1}\left(\frac{\left(\frac{2m+1}{2m+2}\right)^{2m+1}}{2m(2m-1)}\;\frac{\arc\tan\left(\frac{\pi}{\log(2m+1)}\right)}{\pi}-\frac{1}{\pi^2(2m-1)2^{2m+1}}\right)\right.$$

$$\left.-\frac{1}{\pi^2 m}-\frac{(m+1)(2m+1)}{4\pi^2 m(m-1)(2m-1)}\right\}>\frac{1}{8(2m+1)^2},$$

or

$$C^m{}_{2m-1}\left\{\frac{\left(\frac{2m+1}{2m+2}\right)^{2m+1}}{2m(2m-1)}\;\frac{\arc\tan\frac{\pi}{\log(2m+1)}}{\pi}-\frac{1}{\pi^2(2m-1)2^{2m+1}}\right\}$$

$$>\frac{(m+1)(2m+1)}{4\pi^2 m(m-1)(2m-1)}+\frac{1}{\pi^2 m}+\frac{1}{16m},$$

which is fulfilled for $m \geqq 5$.

The inequality $\Sigma A_i(i/n) > 1/8$ is also fulfilled for $n = 7$ by direct calculation, but not for $n = 3$.

9. The coefficients A_{5k} do not alternate in sign. Noting in particular that (A), § 6, is valid,

$$N_0 + 1 \leqq N_1.$$

Since $N_1 + 1 \leqq N_2$, it follows that $N_0 + 2 \leqq N_2$, where N_2 is the number of sign changes of $R_{2m-2}(t)$ as t increases from 0 to 1 for $m = 2$. From tables the graph of $R_{2m-2}(t) = t - \sum_{i/5\leqq t} A_i$ presents precisely three changes of sign on $0 < t < 1$. Hence,

$$N_0 + 2 \leqq 3,$$

or $N_0 = +1$, which was to be shown.

If $n = 3$, there are only two coefficients, which are equal, and hence there is no question of alternation of sign. The expansion has, in this case, the following form:

$$\int_0^1 f(x)dx = \sum_1^2 A_i f(i/3) - \sum_{\nu=1}^{s-1} K_3{}^{2\nu}\frac{\Delta f^{(2\nu-1)}(0)}{(2\nu)!}$$

$$+\int_0^1 \frac{f^{(2s)}(t)}{(2s)!}\sum_1^2 A_i\{\bar{B}_{2s}(i/3-t) - B_{2s}(i/3)\}dt,$$

where

$$K_3{}^{2\nu} = A_1 B_{2\nu}(1/3) + A_2 B_{2\nu}(2/3) = -\frac{(1-3^{1-2\nu})}{2}B_{2\nu},$$

hence, $K_3{}^{2\nu}$ alternates in sign for $\nu = 1, 2, 3, \cdots$.

The mean value theorem can be applied as before to put the remainder in the previously given form since

$$Q_{2s}(t) = \tfrac{1}{2}\{\bar{B}_{2s}(1/3 - t) + \bar{B}_{2s}(2/3 - t) - 2B_{2s}(1/3)\}, \ (s = 1, 2, \cdot \cdot \cdot);$$

does not change sign on $0 < t < 1$. By direct methods this is evident, for on $0 < t < 1/3$,

$$Q_2(t) = t^2,$$

and

$$Q_3(t) = -t^3 + t/6.$$

On $1/3 < t < 2/3$,

$$Q_2(t) = t^2 - t + 1/3,$$
$$Q_3(t) = -t^3 + t/6 + (3/2)(1/3 - t)^2,$$

and thus $\beta_1 = 0$, $\alpha_2 = 1$. From $\beta_s + 1 \leqq \alpha_s$, and $\alpha_s \leqq \alpha_{s-1}$, it is evident that $\beta_2 = \beta_3 = \cdot \cdot \cdot = 0$.

STANFORD UNIVERSITY,
CALIFORNIA.

ON CERTAIN FUNDAMENTAL IDENTITIES DUE TO USPENSKY.*

By W. A. Dwyer.

1. Introduction. In a series of memoirs entitled "Sur les Relations entre les Nombres des Classes des Formes Quadratiques Binaires et Positives,"[1] Uspensky has obtained several very general fundamental formulae involving incomplete numerical functions in three variables. He has made use of these to establish a great variety of interesting and useful arithmetical theorems.[2] Uspensky's proofs of the fundamental formulae are purely arithmetic and of great simplicity, but give no clue as to how a systematic determination of such formulae may be made. They are in the nature of a priori verifications. The structure of the identities suggests that they may be gotten from equivalent identities involving the theta functions by means of the method of paraphrase.[3] If such an identity can be found, it will suggest a systematic determination, by analytical means, of all identities of this type. From this set of identities it would then be possible to pass back, by means of the method of paraphrase, to other general identities, and then to their application to arithmetic. Bell,[4] for example, has discovered a theta function identity which paraphrases into a certain fundamental formula of Uspensky involving complete numerical functions. In this paper we shall establish two of the formulae involving *incomplete* numerical functions as special cases of a general formula which, in turn, results from the paraphrase of a rather peculiar theta identity.

2. Let $F(x, y, z)$ be a function defined for integral values of the arguments and subject to the parity conditions

$$F(-x, -y, -z) = -F(x, y, z), \qquad F(0, 0, 0) = 0.$$

Then there exists an identity involving *incomplete* numerical functions in three variables

* Received January 26, 1937.

[1] *Bulletin de l'Académie des Sciences de l'U. S. S. R.*, 1925, 1926.

[2] J. V. Uspensky, *loc. cit.*, *Quatrième Memoire*, 1926, pp. 547-566. Also, *American Journal of Mathematics*, vol. 50 (1928), pp. 93-122; *Bulletin of American Mathematical Society*, vol. 36 (1930), pp. 743-754.

[3] E. T. Bell, *Transactions of the American Mathematical Society*, vol. 22 (1921), pp. 1-30, and 198-219.

[4] E. T. Bell, *Bulletin of the American Mathematical Society*, vol. 32 (1926), pp. 682-688.

$$\text{I)}\quad \Sigma F(\delta+i,\delta-d+i,i)$$
$$=\Sigma\left\{F\left(\frac{\Delta'+\Delta}{2},\frac{\Delta'-\Delta}{2},\Delta-h\right)-F\left(-h,\frac{\Delta'-\Delta}{2},\Delta-h\right)\right\}$$
$$+e(n)T+a(n)L,$$
$$T=\sum_{j=1}^{s-1}\{F(j,j,s)-F(s,j,s)\},\qquad L=\Sigma F(r,0,r-t),$$

with integral partitions

II) a) $n=i^2+2d\delta,\quad i\gtreqless 0,\quad \delta>0,\quad d>0,$

b) $n=h^2+\Delta\Delta',\quad h\gtreqless 0,\quad 0<\Delta<\Delta',\quad \Delta'\equiv\Delta\ (\mathrm{mod}\ 2),$

c) $n=s^2,\quad s>0,\quad e(n)=0$ or 1 according as n is not or is a perfect square,

d) $n=r^2+t^2,\quad r>0,\quad t>0,\quad a(n)=0$ or 1 according as n is not or is a sum of two squares.

If parity conditions be restricted, corresponding identities result as follows:

$$\text{III)}\quad\begin{cases} F(-x,y,z)=F(x,y,z),\ F(x,-y,-z)\\ \qquad=-F(x,y,z),\ F(x,0,z)=0,\\ \Sigma F(\delta+i,\delta-d+i,i)\\ =\Sigma\left\{F\left(\frac{\Delta'+\Delta}{2},\frac{\Delta'-\Delta}{2},\Delta-h\right)-F\left(h,\frac{\Delta'-\Delta}{2},\Delta-h\right)\right\}+e(n)T,\end{cases}$$

$$\text{IV)}\quad\begin{cases} F(-x,y,z)=-F(x,y,z),F(x,-y,-z)=F(x,y,z),F(x,0,z)=0,\\ \Sigma F(\delta+i,\delta-d+i,i)\\ =\Sigma\left\{F\left(\frac{\Delta'+\Delta}{2},\frac{\Delta'-\Delta}{2},\Delta-h\right)+F\left(h,\frac{\Delta'-\Delta}{2},\Delta-h\right)\right\}+e(n)T,\end{cases}$$

Formulae III and IV are the same as certain formulae discovered by Uspensky[5] and proved by purely arithmetic methods.

3. The theta identity and its paraphrase. Formula I results from the attempt to find a theta-function identity which would paraphrase into III and IV. The procedure consisted in going backwards from these two, and involved the selection of terms, which when arithmetized would meet the conditions II(a) and II(b), and proper adjustment of the arguments. The left side of III or IV suggests the product of a theta and a function of the type

$$\phi_{abc}(x,y)=\frac{\vartheta'_1\vartheta_a(x+y)}{\vartheta_b(x)\vartheta_c(y)}.$$

[5] J. V. Uspensky, *Quatrième Memoire, loc. cit.*, and "On incomplete numerical functions," *Bulletin of the American Mathematical Society, loc. cit.*, p. 746.

The function desired, and its arithmetic equivalent [6] is

$$
\text{V)}\quad \vartheta_3(x+y+z)\phi_{111}(x+y,-y)
\\ = 4\Sigma q^{i^2+2d\delta}\,\Sigma \sin 2[(\delta+i)x+(\delta+i-d)y+iz]
\\ + \{\operatorname{ctn}(x+y)-\operatorname{ctn}(y)\}\cdot\Sigma\, q^{i^2}\cos 2i(x+y+z).
$$

For the terms of III (or IV) involving incomplete numerical functions we shall employ an expression

$$
\text{VI)}\quad \chi(x,y)=\sum_{r=-\infty}^{\infty} q^{r^2}e^{-2iry}\operatorname{ctn}(x-r\pi\tau)=\operatorname{ctn}x+2\sum_{n=1}^{\infty}q^{n^2}\sin 2ny
\\ +4\,\Sigma\, q^n\,\Sigma \sin[(\delta-d)x+2dy],
\\ 0<n=d\delta,\qquad 0<d<\delta,\qquad \delta\equiv d \pmod 2,
$$

which appears as a term in the Fourier development of certain pseudo-periodic functions.[7] By synthesis, we arrived at the relation

$$
\text{VII)}\quad \vartheta_3(x+y+z)\phi_{111}(x+y,-y)
\\ =\vartheta_3(z)\chi(x+y,x+z)-\vartheta_3(x+z)\chi(y,z).
$$

An independent proof of this result will appear in § 4 and we shall proceed with the paraphrase of VII. Applying to it the arithmetized expansions of V and VI and the formula $\vartheta_3(x)=\sum_{n=-\infty}^{\infty}q^{n^2}\cos(2nx)$, we obtain

$$
\text{VII)}\quad 4\,\Sigma\, q^{i^2+2d\delta}\,\Sigma \sin 2[(\delta+i)x+(\delta-d+i)y+iz]
\\ =4\,\Sigma\, q^{h^2+\Delta\Delta'}\{\Sigma \sin[(\Delta'+\Delta)x+(\Delta'-\Delta)y+2(\Delta-h)]
\\ -\Sigma \sin[-2hx+(\Delta'-\Delta)y+2(\Delta-h)z]\}
\\ +\operatorname{ctn}(x+y)\{\Sigma\, q^{i^2}\{\cos(2iz)-\cos 2i(x+y+z)\}\}
\\ -\operatorname{ctn}(y)\{\Sigma\, q^{i^2}\{\cos 2i(x+z)-\cos 2i(x+y+z)\}\}
\\ +2\,\Sigma\, q^{i^2+t^2}\,\Sigma\{\cos(2iz)\sin 2t(x+z)-\cos 2i(x+z)\sin 2tz\},
$$

---

[6] Cf. E. T. Bell, "Theta expansions useful in arithmetic," *Messenger of Mathematics*, vol. 53 (1924), pp. 166-176.

[7] M. A. Basoco, "Fourier developments for certain pseudo-periodic functions in two variables," *American Journal of Mathematics*, vol. 54, no. 2 (1932), p. 242. In this connection we shall exhibit the periodicity properties of $\chi(x,y)$. The function is obviously periodic when x or y is increased by π. Furthermore

$$
\chi(x+n\pi\tau,y+n\pi\tau)=\chi(x,y),
$$

$$
\chi(x,y+n\pi\tau)=e^{-2nix}\chi(x,y)
\\ +ie^{-2nix}\{q^{-n^2}e^{-2ni(y-x)}+2\sum_{k=1}^{n-1}q^{-(n-k)^2}e^{-2i(n-k)(y-x)}+1\}\vartheta_3(y),
$$

$$
\chi(x+n\pi\tau,y)=q^{n^2}e^{2ni(x-y)}\chi(x,y)
\\ -i\{q^{n^2}e^{2ni(x-y)}+2\sum_{k=1}^{n-1}q^{n^2-k^2}e^{2i(n-k)(x-y)}+1\}\vartheta_3(y).
$$

where the n appearing in the $\sum_{n=1}^{\infty}$ term of VI has been replaced by t, and $i, d, \delta, h, \Delta, \Delta', t$ are subject to the conditions II.

In the terms involving cotangents we change the index of summation from i to s (bringing in the multiplier 2, since $i \gtreqless 0$ while $s > 0$), combine the differences of cosines into the product of two sines, and apply the formula, $\sin(au)\operatorname{ctn}(u) = \sum_{k=0}^{a-1} \cos(a-2k)u$. After combining terms and making an obvious change in the index of summation, our expression contributes

$$+ 4e(n)T + 2 \sum q^{t^2}\{\sin 2tz - \sin 2t(x+z)\}.$$

If we split the last term of VIII into a sum corresponding to $i = 0$ and a sum where $i = r$ (r restricted to positive values), we obtain

$$+ 4a(n)L - 2 \sum q^{t^2}\{\sin(2tz) - \sin 2t(x+z)\}.$$

Putting the last two results in VIII, changing all arguments to their half-values, and paraphrasing, we arrive at I.

4. Proof of the theta identity. Consider the left-hand side as a function of y alone. Then

$$f(y) = \vartheta_3(x+y+z)\phi_{111}(x+y, -y) = -\frac{\vartheta'_1\vartheta_3(x+y+z)\vartheta_1(x)}{\vartheta_1(x+y)\vartheta_1(y)},$$
$$f(y+n\pi\tau) = q^{n^2}e^{2ni(y-z)}f(y), \qquad f(y+n\pi) = f(y).$$

The residues at the simple poles, $y = 0 + n\pi\tau$, $y = -x + n\pi\tau$, are respectively $-q^{n^2}e^{-2niz}\vartheta_3(x+z)$ and $+q^{n^2}e^{-2ni(x+z)}\vartheta_3(z)$. Let C represent the contour in the y-complex plane composed of $(n+1)$ cells (of width π) above and n cells below the real axis and consider the auxiliary function $\phi(t) = f(t)/\tan(t-y)$ which has poles at $t = y$, $t = y + \pi$, $t = n\pi\tau$, $t = -x + n\pi\tau$. The residue at $t = y$ is $f(y)$. Derange the mesh so that poles lie within the boundary, apply Cauchy's Theorem to $\phi(t)$ around C, and allow *n to become infinite.*[8] Thus,

$$\begin{aligned}\frac{1}{2\pi i}\int_C \phi(t)\,dt = 0 = \sum \text{Residues} \\ = f(y) + \vartheta_3(x+z)\sum_{n=-\infty}^{\infty} q^{n^2}e^{-2niz}\operatorname{ctn}(y-n\pi\tau) \\ - \vartheta_3(z)\sum_{n=-\infty}^{\infty} q^{n^2}e^{-2ni(x+z)}\operatorname{ctn}(x+y-n\pi\tau).\end{aligned}$$

The last relation is the same as our identity VII.

---

[8] Cf. M. A. Basoco, "Fourier developments for certain pseudo-periodic functions in two variables," *loc. cit.*, pp. 244-245.

In his arithmetic proof of III and IV Uspensky splits the left-hand side into three sums S_1, S_2, S_3 according as $(\delta - d + i)$ is > 0, < 0, or $= 0$, and sets up a pair of transformations establishing a one-to-one correspondence between the solutions of $i^2 + 2d\delta$ and those of $h^2 + \Delta\Delta'$ which obey the restrictions II(a) and II(b). $S_1 + S_2$, together with *stated parity conditions*, gives us identities III and IV. S_3 obviously vanishes because of the condition $F(x, 0, z) = 0$. If, however, we change the parity conditions to agree with those of I, the more general condition $F(0, 0, 0) = 0$ demands that we consider the contribution of the term S_3. From II(a)

$$n = i^2 + 2d\delta. \text{ If } \delta - d + i = 0, \text{ then } n = d^2 + \delta^2.$$

Consequently

$$\Sigma F(\delta + i, \delta + i - d, i)$$

is of the form

$$S_3 = \Sigma F(d, 0, d - \delta),$$

which is of the same form and has the same partitions as the term $a(n)L$ appearing in I.

5. In conclusion, we may point out that the results obtained by Uspensky from his fundamental formulae are implicitly contained in our theta identity. The theta identity suggests other similar products of the theta and ϕ-functions which, when treated in an analogous manner, will lead to general fundamental formulae of the same type as those of Uspensky.

UNIVERSITY OF NEBRASKA.

AN EXTENSION OF BERNSTEIN'S THEOREM ASSOCIATED WITH GENERAL BOUNDARY VALUE PROBLEMS.*

By W. H. McEwen.

Introduction. Consider the n-th order differential system

$$(1)\qquad \begin{aligned} &\frac{d^n u}{dx^n} + P_2(x)\frac{d^{n-2}u}{dx^{n-2}} + \cdots + P_n(x)u + \lambda u = 0, \\ &U_j(u) = 0, \qquad\qquad (j = 1, 2, \cdots, n), \end{aligned}$$

in which the functions $P_2(x), \cdots, P_n(x)$ are continuous and have continuous derivatives of all orders on $a \leqq x \leqq b$, and the U's are n linearly independent conditions involving $u^{(r)}(a), u^{(r)}(b)$, $r = 0, 1, \cdots, n-1$. The general nature of the solutions and the expansion problems connected with this system have been discussed by Birkhoff,[1] Tamarkin,[2] Milne[3] and Stone.[4] Let the characteristic values of the system, taken in the order of magnitude of their moduli, be $\lambda_1, \lambda_2, \cdots$, and let $u_1(x), u_2(x), \cdots$ be the corresponding characteristic solutions. The values λ_k are then the poles of the Green's function of the system. Assume that the boundary conditions are normalized and regular.[5] Assume further that the values λ_k give rise to simple poles of the Green's function when k is large.[6]

Let $S_N(x) = \sum_{j=1}^{N} a_j u_j(x)$ be an arbitrary linear combination of the solutions corresponding to the first N characteristic values, and let L be the maximum value of $|S_N(x)|$ on $a \leqq x \leqq b$. It is the purpose of this paper to establish the following two theorems:

Theorem 1. *On the interval* $a \leqq x \leqq b$

$$|S'_N(x)| \leqq qN^2L,$$

* Received October 12, 1936; revised January 25, 1937.

[1] G. D. Birkhoff, "Boundary value and expansion problems etc.," *Transactions of the American Mathematical Society*, vol. 9 (1908), pp. 373-395.

[2] J. Tamarkin, *Rendiconti del Circolo di Palermo*, vol. 34 (1912), pp. 345-395.

[3] W. E. Milne, *Transactions of the American Mathematical Society*, vol. 19 (1918), pp. 143-156.

[4] M. H. Stone, "A comparison of the Series of Fourier and Birkhoff," *Transactions of the American Mathematical Society*, vol. 28 (1926), pp. 695-761.

[5] For definitions of these terms see Birkhoff, *loc. cit.*, p. 382.

[6] For a discussion of this assumption see footnote 12.

where q is a positive constant independent of N.

THEOREM 2. *On the interval $a+\delta \leqq x \leqq b-\delta$*

$$|S'_N(x)| \leqq QNL,$$

where Q is a positive constant independent of N.

These theorems are analogous respectively to the theorems of Markoff and Bernstein as applied to polynomial sums. In connection with Theorem 2 it should be noted that the limit QNL cannot in general be extended to the end points a and b, as may be shown by an example,[7] although in certain special cases the limit does apply uniformly to the whole interval. Examples of the latter are the systems that give rise to sums of Fourier or Sturm-Liouville type. The Sturm-Liouville case has been treated by Miss E. Carlson,[8] who has also proved Theorems 1 and 2 for the case of a special 3rd order system.[9]

The proofs given here are based on a number of results to be found in Professor Stone's paper.[10] This paper will be referred to hereafter as (S). The writer wishes to acknowledge his indebtedness to Professor Stone for valuable suggestions in connection with the form of presentation of the proofs.

Preliminary discussion. We can assume, without loss in generality, that the interval of x is $0 \leqq x \leqq 1$, and that the maximum value of $|S_N(x)|$ on $(0,1)$ is 1.

Let $G(x,y;\lambda)$ be the Green's function of system (1). The characteristic values of λ are then the poles of G. The facts concerning the nature and distribution of these values are well known. They form two infinite sequences in the complex λ-plane, given asymptotically by the formulas[11]

$$\begin{aligned} \lambda_k' &= -(2k\pi i)^n(1+E'/k), \\ \lambda_k'' &= -(-2k\pi i)^n(1+E''/k), \end{aligned} \tag{2}$$

[7] The system $du/dx+\lambda u=0$, $u(0)+u(1)=0$ gives rise to sums $S_N(x)$ which are sine and cosine sums in the variable $X=\pi x$ which ranges over a part only of a period interval $0 \leqq X \leqq \pi$. It follows then from the well known form of Bernstein's theorem relating to a trigonometric sum on a part of a period interval, that the limit QNL can be assigned only to an interval which is interior to $0 \leqq X \leqq \pi$, and hence to an interval of x which is interior to $(0,1)$.

[8] E. Carlson, *Transactions of the American Mathematical Society*, vol. 26 (1924), pp. 230-240.

[9] E. Carlson, *Transactions of the American Mathematical Society*, vol. 28 (1926), pp. 435-447; pp. 439-447.

[10] M. H. Stone, *loc. cit.*

[11] See Birkhoff, *loc. cit.*, p. 383.

where E', E'' are bounded functions of k. For large values of k, in accordance with the assumption made on page 1,[12] the poles of G are simple. Hence if multiple poles exist they are limited in number. Let C_N be a circle of the λ-plane with centre at the origin which includes within its boundary the first N poles of G, $\lambda_1, \lambda_2, \cdots, \lambda_N$ and no others. Then the sum $S_N(x)$ may be represented identically by the contour integral [13]

$$(3) \qquad S_N(x) = \frac{1}{2\pi i}\int_0^1 S_N(y)\left[\int_{C_N} G(x, y; \lambda)\, d\lambda\right] dy,$$

provided the poles in question are all simple. If, however, certain of these poles are multiple, the integral given above will represent a sum $\sigma_N(x)$ obtained from $S_N(x)$ by replacing the terms corresponding to the multiple poles λ_s by terms of the form $\int_0^1 S_N(y)\, R_s(x, y)\, dy$, where $R_s(x, y)$ is the residue of G at $\lambda = \lambda_s$. But the terms involved in this change are bounded independently of N, and their number also is independent of N. Hence it follows that S'_N and σ'_N are of the same order of magnitude with respect to N. Thus, for the purpose of our discussion, there is no loss in generality in assuming that $S_N(x)$ is represented by (3).

A more useful form of (3) is obtained by placing $\lambda = \rho^n$. Under this transformation the entire λ-plane is made to correspond to a sector Σ in the ρ-plane, composed of two adjacent sectors of the following set of $2n$ equal sectors:

$$S\colon\ l\pi/n \leqq \arg \rho \leqq (l+1)\pi/n, \qquad (l = 1, 2, \cdots, 2n-1).$$

The path of integration will then become the arc Γ which the sector Σ cuts off from the circle with centre at the origin in the ρ-plane and radius equal to the n-th root of the radius of C_N. Hence we can write

$$(4) \qquad S_N(x) = \frac{1}{2\pi i}\int_0^1 S_N(y)\left[\int_\Gamma n\rho^{n-1} G(x, y; \rho^n)\, d\rho\right] dy.$$

---

[12] The assumption referred to is the one which demands that the poles of G be simple when k is large. This condition is not highly restrictive inasmuch as it is automatically satisfied if the system is regular and of odd order, and is in general satisfied if the system is regular and of even order. In the case of even order, however, it may happen that pairs of characteristic values coincide to give double values, and these in turn will give rise to either simple or double poles of G. If the system is self-adjoint the double values give rise to simple poles, but otherwise it is possible to have infinitely many double poles. Tamarkin has given an example of a regular system, $n = 2$, with infinitely many double poles (see Stone and Tamarkin, "Remarks on a paper by Dr. Tautz," *Acta Mathematica*, 1931). It is this type of system which our hypothesis rules out.

[13] See Birkhoff, *loc. cit.*, p. 379.

A special case of (1) is the system (discussed in S, pp. 709-711)

$$d^n u/dx^n + \lambda u = 0, \quad u^{(j)}(0) - u^{(j)}(1) = 0, \qquad (j = 0, 1, \cdots, n-1),$$

which gives rise to sums of Fourier type. Let $\bar{G}(x, y; \lambda)$ denote the Green's function in this case. The arc Γ may be drawn so as to avoid the poles of $\bar{G}$ as well as those of G. Then the integral

$$(5) \qquad \frac{1}{2\pi i}\int_0^1 S_N(y)\left[\int_\Gamma n\rho^{n-1}\bar{G}(x, y; \rho^n)\,d\rho\right]dy$$

defines a sum $T_N(x)$ which is a trigonometric sum of order $\bar{N}/2$ on the period interval $0 \leqq x \leqq 1$, where $|\bar{N} - N| \leqq K$ independent of the radius of Γ. The latter relation means that $O(\bar{N}/2) = O(N)$.

In regard to the arc Γ we shall demand further that it be kept uniformly away from the poles of G and $\bar{G}$ when the radius is large.

We next define a set of constants which play an important rôle in the asymptotic formulas for G, $\bar{G}$ and their derivatives. These are the n n-th roots of -1, denoted by $\omega_1, \omega_2, \cdots, \omega_n$. For values of ρ on any given sector S let the subscripts be so chosen that

$$R(\rho\omega_1) \leqq R(\rho\omega_2) \leqq \cdots \leqq R(\rho\omega_n), \qquad R \equiv \text{“the real part of.”}$$

Then, if the system is of odd order

$$n = 2\mu - 1,$$

$R(\rho\omega_\mu) = 0$ on the bisecting ray of S, so that in one half of S $R(\rho\omega_\mu) < 0$ whereas in the other half $R(\rho\omega_\mu) > 0$. Let these two halves be denoted by S' and S'' respectively. Thus we have

$$R(\rho\omega_1) \leqq \cdots \leqq R(\rho\omega_\mu) \leqq 0 \leqq R(\rho\omega_{\mu+1}) \leqq \cdots \leqq R(\rho\omega_n) \quad \text{on} \quad S',$$
$$R(\rho\omega_1) \leqq \cdots \leqq R(\rho\omega_{\mu-1}) \leqq 0 \leqq R(\rho\omega_\mu) \leqq \cdots \leqq R(\rho\omega_n) \quad \text{on} \quad S''.$$

On the other hand, if the system is of even order

$$n = 2\mu,$$

$\omega_\mu = -\omega_{\mu+1}$ and on one of the bounding rays of S $R(\rho\omega_\mu) = R(\rho\omega_{\mu+1}) = 0$, so that, throughout the whole of S,

$$R(\rho\omega_1) \leqq \cdots \leqq R(\rho\omega_\mu) \leqq 0 \leqq R(\rho\omega_{\mu+1}) \leqq \cdots \leqq R(\rho\omega_n).$$

These results enable us to state the conditions under which the exponential functions $e^{\rho\omega_j(x-y)}$, $j = 1, 2, \cdots, n$, occurring in the asymptotic formulas for G, $\bar{G}$, etc. are bounded in the form

$$|\, e^{\rho\omega_j(x-y)} \,| \leqq 1, \qquad 0 \leqq x, y \leqq 1, \tag{6}$$

for all values of ρ in question. This inequality holds whenever the real part of $\rho\omega_j(x-y)$ is negative or zero, and hence the specific requirements to be met are as follows:

Case 1. $n = 2\mu - 1$.

$$j \begin{cases} = 1, \cdots, \mu; & x \geqq y; \quad \rho \text{ on } S', \\ = \mu + 1, \cdots, n; & x \leqq y; \quad \rho \text{ on } S'', \\ = 1, \cdots, \mu - 1; & x \geqq y; \quad \rho \text{ on } S'', \\ = \mu, \cdots, n; & x \leqq y; \quad \rho \text{ on } S''. \end{cases} \tag{6'}$$

Case 2. $n = 2\mu$.

$$j \begin{cases} = 1, \cdots, \mu; & x \geqq y; \quad \rho \text{ on } S, \\ = \mu + 1, \cdots, n; & x \leqq y; \quad \rho \text{ on } S. \end{cases} \tag{6''}$$

In the explicit formulas for G, $\bar{G}$, etc., which will be used presently, it will be seen that these conditions are satisfied in every instance, so that the exponentials occurring in these formulas are bounded in the manner of (6).

We now observe a number of lemmas, the essential parts of the proofs of which are based on results found in (S). The notation $\{A; B\}$ is used to indicate that A is to be taken if $x \geqq y$, and B if $x \leqq y$. The letter R is used to denote the radius of the circular arc Γ, so that for values of ρ on Γ $|\rho| = R$. In outlining the proofs it is necessary to treat separately the cases $n = 2\mu - 1$ and $n = 2\mu$, inasmuch as the asymptotic formulas concerned are different in these two cases. It is sufficient, however, to treat only one of the two equal sectors S which make up Σ. The part of Γ belonging to S will be denoted by γ, and the two halves of it corresponding to S', S'' by γ', γ'' respectively.

Lemma 1. *For values of* ρ *on* Γ

$$n\rho^{n-1}\left\{\frac{\partial G}{\partial x}; \frac{\partial G}{\partial x}\right\} = O(R)$$

uniformly on $0 \leqq x$, $y \leqq 1$.

Case 1. $n = 2\mu - 1$. For values of ρ on γ' the following asymptotic formula for the expression in question is given in (S, p. 745, with $k = 1$):

$$\begin{aligned} n\rho^{n-1}&\left\{\frac{\partial G}{\partial x}; \frac{\partial G}{\partial x}\right\} \\ &= \left\{ F_{11}{}^0 - \sum_{j=1}^{\mu} e^{\rho\omega_j(x-y)}\,\frac{m_j(x,y,\rho)}{\rho};\ F_{11}{}^1 + \sum_{j=\mu+1}^{n} e^{\rho\omega_j(x-y)}\frac{m_j(x,y,\rho)}{\rho} \right\} \\ &\qquad + \frac{\Delta_1{}^{(1)}}{[\theta_0] + e^{\rho\omega_\mu}[\theta_1]}. \end{aligned} \tag{7}$$

where

$$F_{11}{}^0 \equiv -\sum_{j=1}^{\mu} \omega_j e^{\rho\omega_j(x-y)}(A_{11}(x) + B_1(y) + \rho\omega_j),$$

$$F_{11}{}^1 \equiv +\sum_{j=\mu+1}^{n} \omega_j e^{\rho\omega_j(x-y)}(A_{11}(x) + B_1(y) + \rho\omega_j),$$

$$([\theta_0] + e^{\rho\omega_\mu}[\theta_1])^{-1} = O(1),^{14}$$

and $\Delta_1{}^{(1)}$ is a determinant of order $n+1$, which, if expanded according to the elements of the first row, has the form

$$\Delta_1{}^{(1)} \equiv \sum_{j=1}^{\mu} (\rho\omega_j) e^{\rho\omega_j x} Q_j(x,y,\rho) + \sum_{j=\mu+1}^{n} (\rho\omega_j) e^{\rho\omega_j(x-1)} Q_j(x,y,\rho).$$

The functions $m_j(x,y,\rho)$, $A_{11}(x)$, $B_1(y)$, $Q_j(x,y,\rho)$ are bounded in their respective variables on $0 \leqq x, y \leqq 1$ when R is large. A similar formula holds when ρ is on γ'', the only difference being that the summations are extended over the ranges $(1, \mu - 1)$ and (μ, n).

On examining the individual terms in these expressions it is seen that conditions (6′) are satisfied, so that the exponentials are all bounded in the manner of (6). Hence the terms are either $O(1)$ or $O(R)$ for values of ρ on γ' and γ''; that is, $n\rho^{n-1}\left\{\dfrac{\partial G}{\partial x}; \dfrac{\partial G}{\partial x}\right\} = O(R)$ on γ.

Case 2. $n = 2\mu$. The asymptotic formula for $n\rho^{n-1}\left\{\dfrac{\partial G}{\partial x}; \dfrac{\partial G}{\partial x}\right\}$ when ρ is on γ is similar in form to the one used in Case 1, ρ on S'; an explicit expression for it is given in (S, p. 760, with $k = 1$). The exponentials involved satisfy conditions (6″) and so are bounded as in (6), and the steps of the proof go through exactly as in Case 1.

LEMMA 2. $\displaystyle\int_\Gamma n\rho^{n-1}[G(x,y;\rho^n) - G(x,y;\rho^n)]d\rho = O(1)$ uniformly on $0 < \delta \leqq x \leqq 1 - \delta$, $0 \leqq y \leqq 1$.

Case 1. $n = 2\mu - 1$. This is (S, Theorem VII, p. 716, with $l = 0$).

Case 2. $n = 2\mu$. The proof in this case is analogous to that of the theorem cited above. In brief outline it is as follows:

For ρ on γ we have, by (S, p. 755),

---

14 The arc Γ must be kept uniformly away from the poles of G when R is large, so as to make this fraction bounded for large values of R. In view of the manner of distribution of the characteristic values it is clear that this can always be done.

$$n\rho^{n-1}G(x,y;\rho^n) \equiv \{-\sum_{j=1}^{\mu} e^{\rho\omega_j(x-y)}[\omega_j];+\sum_{j=\mu+1}^{n} e^{\rho\omega_j(x-y)}[\omega_j]\}\text{ [15]}$$
$$+\frac{\Delta_3}{[\theta_1]e^{2\rho\omega_\mu}+[\theta_0]e^{\rho\omega_\mu}+[\theta_2]}.$$

The arc γ being kept uniformly away from the poles of G when R is large, the denominator of the second term is bounded away from zero. The numerator of the second term is a determinant of order $n+1$, which, if expanded according to the elements of the first row, has the form of a linear combination of the functions $e^{\rho\omega_j x}$ $(j=1,\cdots,\mu)$, $e^{\rho\omega_j(x-1)}$ $(j=\mu+1,\cdots,n)$ with coefficients which are bounded functions of x, y, ρ. The second term above may thus be written in the form

$$\sum_{j=1}^{\mu} e^{\rho\omega_j x}M_j(x,y,\rho)+\sum_{j=\mu+1}^{n} e^{\rho\omega_j(x-1)}M_j(x,y,\rho),$$

in which the functions $M_j(x,y,\rho)$ are uniformly bounded on $0\leqq x,y\leqq 1$ for large values of R.

In the special case of system (1) which defines $\bar{G}$, a corresponding formula will represent $n\rho^{n-1}\bar{G}$ on γ. It will be identical with the one given above except for different terms of higher order in the asymptotic forms $[\omega_i]$ and different functions $\bar{M}_i(x,y,\rho)$. Hence on subtracting these results we obtain

$$n\rho^{n-1}(G-\bar{G})\equiv\left\{-\sum_{j=1}^{\mu} e^{\rho\omega_j(x-y)}\frac{a_j}{\rho};+\sum_{j=\mu+1}^{n} e^{\rho\omega_j(x-y)}\frac{a_j}{\rho}\right\}$$
$$+\sum_{j=1}^{\mu} e^{\rho\omega_j x}(M_j-\bar{M}_j)+\sum_{j=\mu+1}^{n} e^{\rho\omega_j(x-1)}(M_j-\bar{M}_j).$$

The exponentials in { } are bounded as in (6). Hence $\int_\gamma n\rho^{n-1}(G-\bar{G})\,d\rho$ is expressible in terms of integrals of the form

$$\int_\gamma\frac{m\,d\rho}{\rho},\ \int_\gamma e^{\rho\omega_j x}m\,d\rho\ (j=1,\cdots,\mu),\ \int_\gamma e^{\rho\omega_j(x-1)}m\,d\rho\ (j=\mu+1,\cdots,n),$$

which, according to (S, Lemmas III, IV′, V′, p. 714 and pp. 754-755, with $k=l=0$) are uniformly bounded on $0<\delta\leqq x\leqq 1-\delta$.

LEMMA 3. $\int_\Gamma n\rho^{n-1}\left\{\frac{\partial G}{\partial x}-\frac{\partial\bar{G}}{\partial x};\frac{\partial G}{\partial x}-\frac{\partial\bar{G}}{\partial x}\right\}d\rho=O(R)$ *uniformly on* $0<\delta\leqq x\leqq 1-\delta$, $0\leqq y\leqq 1$.

[15] The notation $[W]$ indicates an asymptotic form in ρ in which W is the leading term.

Case 1. $n=2\mu-1$. For ρ on γ' we use again formula (7) of Lemma 1, writing $E_j(x,y,\rho)$ for $Q_j(x,y,\rho)/([\theta_0]+e^{\rho\omega\mu}[\theta_1])$.

$$n\rho^{n-1}\left\{\frac{\partial G}{\partial x};\frac{\partial G}{\partial x}\right\}=\left\{F_{11}{}^0-\sum_{j=1}^{\mu}e^{\rho\omega_j(x-y)}\frac{m_j}{\rho};\ F_{11}{}^1+\sum_{j=\mu+1}^{n}e^{\rho\omega_j(x-y)}\frac{m_j}{\rho}\right\}$$
$$+\sum_{j=1}^{\mu}(\rho\omega_j)e^{\rho\omega_jx}E_j+\sum_{j=\mu+1}^{n}(\rho\omega_j)e^{\rho\omega_j(x-1)}E_j.$$

A similar form holds for $\bar{G}$, with $F_{11}{}^0$, $F_{11}{}^1$, m_j, E_j replaced by $\bar{F}_{11}{}^0$, $\bar{F}_{11}{}^1$, $\bar{m}_j$, $\bar{E}_j$ respectively. Hence on forming the difference of these two formuals we obtain

$$n\rho^{n-1}\left\{\frac{\partial G}{\partial x}-\frac{\partial\bar{G}}{\partial x};\frac{\partial G}{\partial x}-\frac{\partial\bar{G}}{\partial x}\right\}$$
$$\equiv\left\{(F_{11}{}^0-\bar{F}_{11}{}^0)-\sum_{j=1}^{\mu}e^{\rho\omega_j(x-y)}\frac{(m_j-\bar{m}_j)}{\rho}\right.$$
$$\left.;\ (F_{11}{}^1-\bar{F}_{11}{}^1)+\sum_{j=\mu+1}^{n}e^{\rho\omega_j(x-y)}\frac{(m_j-\bar{m}_j)}{\rho}\right\}$$
$$+\sum_{j=1}^{\mu}(\rho\omega_j)e^{\rho\omega_jx}(E_j-\bar{E}_j)+\sum_{j=\mu+1}^{n}(\rho\omega_j)e^{\rho\omega_j(x-1)}(E_j-\bar{E}_j).$$

But

$$F_{11}{}^0\equiv-\sum_{j=1}^{\mu}\omega_je^{\rho\omega_j(x-y)}(A_{11}(x)+B_1(y)+\rho\omega_j),$$

whereas, according to (S, p. 746, with $s=1$),

$$\bar{F}_{11}{}^0\equiv-\sum_{j=1}^{\mu}\omega_je^{\rho\omega_j(x-y)}(\rho\omega_j).$$

Hence, by (6),

$$F_{11}{}^0-\bar{F}_{11}{}^0\equiv-\sum_{j=1}^{\mu}\omega_je^{\rho\omega_j(x-y)}(A_{11}(x)+B_1(y))=O(1).$$

Likewise $F_{11}{}^1-\bar{F}_{11}{}^1=O(1)$. The remaining terms of { } are similarly bounded (approaching zero as $R\to\infty$). Hence the integral of the expression in { } taken over the arc γ' will be of the order R. Moreover the integrals

$$\int_{\gamma'}\rho e^{\rho\omega_jx}(E_j-\bar{E}_j)d\rho \qquad (j=1,\cdots,\mu-1),$$
$$\int_{\gamma'}\rho e^{\rho\omega_j(x-1)}(E_j-\bar{E}_j)d\rho \qquad (j=\mu+1,\cdots,n)$$

converge to zero uniformly on $0<\delta\leqq x\leqq 1-\delta$ as $R\to\infty$, in accordance with (S, Lemma IV, p. 714, with $k=1$). Hence

$$\int_{\gamma'}n\rho^{n-1}\left\{\frac{\partial G}{\partial x}-\frac{\partial\bar{G}}{\partial x};\frac{\partial G}{\partial x}-\frac{\partial\bar{G}}{\partial x}\right\}d\rho=O(R)+\omega_\mu\int_{\gamma'}\rho e^{\rho\omega_\mu x}(E_\mu-\bar{E}_\mu)d\rho.$$

In a like manner we find that the integral of this expression over γ'' is of the form

$$O(R) + \omega_\mu \int_{\gamma''} \rho e^{\rho\omega_\mu(x-1)} (E'_\mu - \bar{E}'_\mu)\, d\rho.$$

It remains now for us to determine the order of the second terms in the two expressions written above. The integrals involved may be put in the form

$$R\int_{\gamma'} e^{\rho\omega_\mu x} m'\, d\rho, \qquad R\int_{\gamma''} e^{\rho\omega_\mu(x-1)} m''\, d\rho,$$

where $m' = \rho(E_\mu - \bar{E}_\mu)/R = O(1)$ on γ', and $m'' = \rho(E'_\mu - \bar{E}'_\mu)/R = O(1)$ on γ''. An application of (S, Lemma V, p. 714, with $k = l = 0$) will then show that the multipliers of R are uniformly bounded on $0 < \delta \leqq x \leqq 1 - \delta$, and hence we conclude that the second terms also are of order R in this interval.

Case 2. $n = 2\mu$. The argument in this case, based on appropriate formulas found in (S), is entirely similar to the one just given.

LEMMA 4. *The number of poles of $G(x, y; \rho^n)$ on the sector Σ enclosed by the arc Γ is given asymptotically by*

$$N \sim \frac{1}{\pi} R,$$

(each pole being counted according to its multiplicity).

This is immediately evident from formulas (2) which give the distribution of the characteristic values of $\lambda = \rho^n$. From this lemma it follows that

$$O(R) = O(N).$$

The proofs of Theorems 1 and 2. On differentiating with respect to x in formula (4) we get

$$S'_N(x) = \frac{1}{2\pi i}\int_0^1 S_N(y) \left[\int_\Gamma n\rho^{n-1} \left\{ \frac{\partial G}{\partial x}; \frac{\partial G}{\partial x} \right\} d\rho \right] dy.$$

But $|S_N(x)| \leqq 1$ on $0 \leqq y \leqq 1$. Hence, by Lemmas 1 and 4,

$$S'_N(x) = \frac{1}{2\pi i}\int_0^1 S_N(y) \left[\int_\Gamma O(R)\, d\rho \right] dy = O(R^2) = O(N^2)$$

on $0 \leqq x \leqq 1$. This proves Theorem 1.

Next, let us consider the trigonometric sum $T_N(x)$ defined by (5). On subtracting it from $S_N(x)$ we have, by reason of Lemma 2,

$$S_N(x) - T_N(x) = \frac{1}{2\pi i}\int_0^1 S_N(y)\left[\int_\Gamma n\rho^{n-1}(G - \bar{G})\,d\rho\right]dy$$
$$= \frac{1}{2\pi i}\int_0^1 S_N(y)O(1)\,dy = O(1)$$

on $0 < \delta \leqq x \leqq 1 - \delta$. But $|S_N(x)| \leqq 1$, and hence

$$T_N(x) = O(1)$$

on the interval $0 < \delta \leqq x \leqq 1 - \delta$, which is interior to the period interval $(0, 1)$. It follows then, from a special form of Bernstein's theorem given by D. Jackson,[16] that

$$T'_N(x) = O(\bar{N}/2) = O(N) \tag{8}$$

uniformly on this interior interval.

Finally, on differentiating with respect to x in the formula for $S_N(x) - T_N(x)$ we obtain, by the help of Lemmas 3 and 4,

$$S'_N(x) - T'_N(x) = \frac{1}{2\pi i}\int_0^1 S_N(y)\left[\int_\Gamma n\rho^{n-1}\left\{\frac{\partial G}{\partial x} - \frac{\partial \bar{G}}{\partial x};\ \frac{\partial G}{\partial x} - \frac{\partial \bar{G}}{\partial x}\right\}d\rho\right]dy$$
$$= \frac{1}{2\pi i}\int_0^1 S_N(y)O(R)\,dy = O(R) = O(N)$$

uniformly on $0 < \delta \leqq x \leqq 1 - \delta$. Hence, by (8),

$$S'_N(x) = T'_N(x) + O(N) = O(N)$$

uniformly on $0 < \delta \leqq x \leqq 1 - \delta$. This proves Theorem 2.

Application to a problem of best approximation. Let $f(x)$ be a given function continuous on $a \leqq x \leqq b$, and let it be required to define for each positive integral value of N a function of best approximation to $f(x)$ of the form

$$S_N(x) = \sum_{k=1}^{N} a_k u_k(x),$$

in which the u's are the characteristic solutions of system (1). This may be done by adopting as a measure of approximation the integral

$$\int_a^b |f(x) - S_N(x)|^r\,dx,$$

where r is any given real constant > 0, and requiring that the coefficients of

[16] D. Jackson, *Transactions of the American Mathematical Society*, vol. 26 (1924), pp. 133-154; p. 145.

$S_N(x)$ be chosen in such a way as to give a minimum value to this integral. It is well known that such determinations of the coefficients can always be made, and when $r > 1$ the result is unique.

The question of the convergence of $S_N(x)$ to $f(x)$ as N becomes infinite may be investigated by methods similar to those used by Jackson [17] in the study of the corresponding problems relating to trigonometric sums and polynomials. These methods involve in an essential way the use of Theorems 1 and 2, and lead to the following general theorem:

THEOREM 3. *If $\pi_N(x)$ be an arbitrary sum of the u's of order N and h_N be the maximum value of $|f(x) - \pi_N(x)|$ on $a \leqq x \leqq b$, then there will exist positive constants C_1 and C_2 independent of N such that*

(a) *on* $a \leqq x \leqq b$, $|f(x) - S_N(x)| \leqq C_1 N^{2/r} h_N$,

(b) *on* $a + \delta \leqq x \leqq b - \delta$, $|f(x) - S_N(x)| \leqq C_2 N^{1/r} h_N$.

Thus the question of convergence is made to depend directly on the degree of approximation represented by h_N, that is on the degree of approximation to $f(x)$ that is possible by sums of the form $\pi_N(x)$. In this connection we have the theorems on the degree of convergence of Birkoff's series given by Milne [18] which enable us to state explicit hypotheses under which the quantities $N^{2/r}h_N$ and $N^{1/r}h_N$ will converge to zero as N becomes infinite. The following theorem is given as typical of what can be done in this direction:

THEOREM. *In the case $r > 1$, if $f(x)$ has a first derivative of limited variation on $a \leqq x \leqq b$, and if $f(x)$ vanishes at a and b, then $h_N = O(1/N)$ so that $S_N(x)$ converges uniformly to $f(x)$ on the sub-interval $a + \delta \leqq x \leqq b - \delta$ as N becomes infinite.*

MOUNT ALLISON UNIVERSITY,
SACKVILLE, N. B., CANADA.

[17] D. Jackson, *American Mathematical Society, Colloquium Publications*, vol. 11, pp. 92-101; see also D. Jackson, *Bulletin of the American Mathematical Society*, December, 1933, pp. 889-906.

[18] W. E. Milne, *loc. cit.*, pp. 154-156.

ABSTRACT COVARIANT VECTOR FIELDS IN A GENERAL ABSOLUTE CALCULUS.*

By A. D. Michal.

Introduction. The elements of a general absolute differential calculus based on a linear connection and the notion of a contravariant vector field only has recently been considered by me.[1] In the present paper additional postulates on the transformation of Banach "coördinates" are considered and a linear connection of *covariant type* is postulated. A brief treatment is then given of a general absolute differential calculus based on covariant vector fields as well as on contravariant vector fields. The ideas centering around the various adjoints of the Fréchet differentials[2] are of fundamental importance here as well as in the instances in which the Banach space is an infinitely dimensional function space.

1. Abstract coördinate transformations. Let E be a Banach space in which there exists a function $[x, y]$ with the following properties:[3]

(1) $[x, y]$ is a bilinear function on E^2 to the real numbers

(2) $[x, y] = [y, x]$

(3) $[x, y]$ is positive definite; i. e., $[x, x] \geqq 0$ and $[x, x] = 0$ if and only if $x = 0$.

Definition. *A function $T^*(\xi)$ on E to E will be said to be the adjoint of a linear function $T(\xi)$ on E to E if*

(1) *$T^*(\xi)$ is a linear function*

(2) $[T(\xi), \eta] = [\xi, T^*(\eta)]$.

Let U_0 be a fixed Hausdorff neighborhood of a Hausdorff[4] space T. We

---

* Presented to the American Mathematical Society under a different title, April, 1936. Received by the Editors February 16, 1937.

[1] Michal (1).

[2] By the notation $f(x, y_1, \cdots, y_n; \lambda)$ we shall always mean the partial Fréchet differential of $f(x, y_1, \cdots, y_n)$ in x with increment λ. Occasionally we shall write $\delta\phi(x)$ for $\phi(x; \delta x)$ and $\underset{\delta x}{\overset{x}{dy}} f(x, y)$ for the partial Fréchet differential of $f(x, y)$ in y evaluated at $y = x$.

[3] For motions and rotations in such spaces see Michal, Highberg and Taylor (1).

[4] More generally one can take a *Fréchet neighborhood space V* and require the coördinate systems to be merely reciprocal (1-1) transformations.

shall assume that there exists an open set $S \subset E$ that is a homeomorphic map of U_0. We postulate the existence of coördinate systems $x(P)$: homeomorphic correspondences mapping Hausdorff neighborhoods onto open sets $\Sigma \subset S$. Suppose $x(P)$ and $\bar{x}(P)$ are coördinate systems for two intersecting Hausdorff neighborhoods U_1 and U_2 respectively and let Σ_1 and Σ_2 be the respective maps in S. Then the intersection of U_1 and U_2 induces a homeomorphic mapping, called a coördinate transformation, of an open subset S_1 of Σ_1 onto an open subset S_2 of Σ_2. We shall denote this coördinate transformation by $\bar{x} = \bar{x}(x)$.

It is convenient to call the Hausdorff neighborhood and the map $\Sigma \subset S$, the geometrical domain and the coördinate domain respectively of the coördinate system. We shall assume that each coördinate transformation $\bar{x}(x)$ and its inverse $x(\bar{x})$ have Fréchet differentials $\bar{x}(x;\delta x)$ and $x(\bar{x};\delta\bar{x})$ throughout the sub-coördinate domains S_1 and S_2 respectively of the coördinate systems $x(P)$ and $\bar{x}(P)$. We shall further assume that $\bar{x}(x;\delta x)$ possesses an adjoint $\bar{x}^*(x;\delta x)$ and that $x(\bar{x};\delta\bar{x})$ has an adjoint $x^*(\bar{x};\delta\bar{x})$. It can be shown readily that $\bar{x}(x;\delta x)$ is a solvable linear function of δx with $x(\bar{x};\delta\bar{x})$ as inverse. From the postulates for $[x, y]$ and the following evident steps

$$\begin{aligned}[\delta\bar{x}, \xi] = [\bar{x}(x; x(\bar{x};\delta\bar{x})), \xi] &= [x(\bar{x};\delta\bar{x}), \bar{x}^*(x;\xi)] \\ &= [\delta\bar{x}, x^*(\bar{x};\bar{x}^*(x;\xi))]\end{aligned}$$

it follows that

$$x^*(\bar{x};\bar{x}^*(x;\xi)) = \xi \quad \text{for all } \xi\epsilon E. \tag{1.1}$$

Similarly

$$\bar{x}^*(x;x^*(\bar{x};\xi)) = \xi \quad \text{for all } \xi\epsilon E. \tag{1.2}$$

From these two results it follows that $\bar{x}^*(x;\delta x)$ is a solvable linear function of δx with $x^*(\bar{x};\delta\bar{x})$ as inverse.

2. Covariant differential of a covariant vector field. The absolute calculus of *contravariant* vector fields has been studied elsewhere.[5] The components of a geometric object have a characteristic law of transformation in the intersection of two Hausdorff neighborhoods. The law of transformation for a contravariant vector field is

$$\bar{\xi}(\bar{x}) = \bar{x}(x;\xi(x)). \tag{2.1}$$

DEFINITION 2.1. *A covariant vector field is a geometric object whose components transform in the intersection of two Hausdorff neighborhoods according to the law*

$$\bar{\eta}(\bar{x}) = x^*(\bar{x};\eta(x)) \tag{2.2}$$

under a transformation of coördinates $\bar{x} = \bar{x}(x)$.

---

[5] Michal (1).

The inverse of (2.2) is

$$\eta(x) = \bar{x}^*(x; \bar{\eta}(\bar{x})). \tag{2.3}$$

In addition to the restrictions of § 1 we shall now assume that the second Fréchet differential $\bar{x}(x; \delta_1 x; \delta_2 x)$ exists continuous in x and that the Fréchet differential

$$\underset{\delta \bar{x}}{\overset{\bar{x}}{d_\mu}} x^*(\sigma; \eta)$$

exists continuous in $\bar{x}$. It can be shown with the aid of theorems proved elsewhere[6] that $x(\bar{x}; \delta_1\bar{x}; \delta_2\bar{x})$ exists continuous in $\bar{x}$. It can also be shown that the adjoint $x^*(\bar{x}; \mu; \lambda)$ of the linear function $x(\bar{x}; \mu; \lambda)$ of λ exists continuous in $\bar{x}$ and that the adjoint $x^*(\bar{x}, \lambda, \mu)$ of the linear function $\underset{\mu}{\overset{\bar{x}}{d_\sigma}} x^*(\sigma; \lambda)$ of μ exists continuous in $\bar{x}$.

THEOREM 2.1. *Under the restrictions on coördinate transformations just described, the following relations are valid*

$$\underset{\mu}{\overset{\bar{x}}{d_\sigma}} x^*(\sigma; \lambda) = x^*(\bar{x}; \mu; \lambda) \tag{2.4}$$

$$\underset{\mu}{\overset{\bar{x}}{d_\sigma}} x^*(\sigma; \lambda) = x^*(\bar{x}, \lambda, \mu). \tag{2.5}$$

The functions $x^*(\bar{x}; \mu; \lambda)$ and $x^*(\bar{x}, \lambda, \mu)$ are bilinear in λ and μ and self adjoint as linear functions of μ.

Proof. On differentiating

$$[x(\bar{x}; \nu), \lambda] = [\nu, x^*(\bar{x}; \lambda)] \tag{2.6}$$

we obtain

$$[x(\bar{x}; \nu; \mu), \lambda] = [\nu, \underset{\mu}{\overset{\bar{x}}{d_\sigma}} x^*(\sigma; \lambda)]. \tag{2.7}$$

Clearly

$$[x(\bar{x}; \nu; \mu), \lambda] = [\nu, x^*(\bar{x}; \mu; \lambda)]. \tag{2.8}$$

Hence from (2.7), (2.8) and the positive definiteness of the inner product there results (2.4).

From (2.7) and the definition of $x^*(\bar{x}, \lambda, \nu)$ we have

$$[x(\bar{x}; \nu; \mu), \lambda] = [\mu, x^*(\bar{x}, \lambda, \nu)]. \tag{2.9}$$

But $x(\bar{x}; \nu; \mu)$ is symmetric in ν and μ so that (2.9) makes clear that

[6] Michal (2); Michal and Elconin (1).

$x^*(\bar{x}, \lambda, \nu)$ is self adjoint as a linear function of ν. Hence (2.5) is valid. Finally the bilinearity of $x^*(\bar{x}; \mu; \lambda)$ and $x^*(\bar{x}, \lambda, \mu)$ in λ and μ follows from the continuity of $\underset{\mu}{\overset{\bar{x}}{d_\sigma}} x^*(\sigma; \lambda)$ in $\bar{x}$ and a theorem of Banach on the linearity of the limit of a sequence of linear functions.[7]

THEOREM 2.2. *A necessary and sufficient condition that* $[\xi(x), \eta(x)]$ *be a scalar invariant for an arbitrary contravariant vector*[8] $\xi(x)$ *is that* $\eta(x)$ *be a covariant vector.*[8]

DEFINITION 2.2. *A covariant linear connection is a geometric object with components* $L(x, \eta(x), \xi(x))$ *that are bilinear functions of a covariant vector* $\eta(x)$ *and a contravariant vector* $\xi(x)$ *and such that in the intersection of the geometrical domains of two coördinate systems* $x(P)$ *and* $\bar{x}(P)$, *the components have the law of transformation*

$$\bar{L}(\bar{x}, \bar{\eta}(\bar{x}), \bar{\xi}(\bar{x})) = x^*(\bar{x}; L(x, \eta(x), \xi(x)) + x^*(\bar{x}; \bar{\xi}(\bar{x}); \eta(x)) \tag{2.10}$$

under the transformation of coördinates $\bar{x} = \bar{x}(x)$.

Let the covariant vector $\eta(x)$ have a continuous differential $\eta(x; \delta x)$. Then from known theorems on Fréchet differentials and from Theorem 2.1 we obtain

$$\bar{\eta}(\bar{x}; \delta\bar{x}) = x^*(\bar{x}; \eta(x; \delta x)) + x^*(\bar{x}; \delta\bar{x}; \eta(x)). \tag{2.11}$$

Hence with the aid of (2.10) we obtain

$$\bar{\eta}(\bar{x}; \delta\bar{x}) - \bar{L}(\bar{x}, \bar{\eta}(\bar{x}), \delta\bar{x}) = x^*\big(\bar{x}; \eta(x; \delta x) - L(x, \eta(x), \delta x)\big). \tag{2.12}$$

The steps are reversible so that we have proved the

THEOREM 2.3. *A necessary and sufficient condition that*

$$\eta(x; \delta x) - L(x, \eta(x), \delta x) \tag{2.13}$$

be a covariant vector whenever $\eta(x)$ *is any continuously differentiable covariant vector is that* $L(x, \eta(x), \delta x)$ *be a covariant linear connection.*

---

[7] Michal (2).

[8] We use contravariant vector and covariant vector as abbreviations for contravariant vector field and covariant vector field respectively. One can, however, define a covariant vector (strict) and contravariant vector (differential) as usual and recast the definitions and theorems (except Theorem 2.3) in the obvious way by substituting contravariant (covariant) vectors for contravariant (covariant) vector fields in the arguments of the linear connections and multilinear forms.

DEFINITION 2.3. *If $L(x, \eta(x), \delta x)$ is a covariant linear connection, then the linear form $\eta(x/\delta x)$ in δx defined by*

$$(2.14) \qquad \eta(x/\delta x) = \eta(x;\delta x) - L(x, \eta(x), \delta x)$$

will be called the covariant differential (based on L) of the covariant vector $\eta(x)$.

Let $\Gamma(x, \xi_1(x), \xi_2(x))$ (not necessarily symmetric) be a linear connection,[9] where $\xi_1(x), \xi_2(x)$ are contravariant vectors. This is to be distinguished from the covariant linear connection of Theorem 2.3. The law of transformation [10] for a linear connection is

$$(2.15) \quad \bar{\Gamma}(\bar{x}, \bar{\xi}_1(\bar{x}), \bar{\xi}_2(\bar{x})) = \bar{x}(x; \Gamma(x, \xi_1(x), \xi_2(x))) + \bar{x}(x; x(\bar{x}; \bar{\xi}_1(\bar{x}); \bar{\xi}_2(\bar{x}))).$$

An equivalent law of transformation to (2.15) is

$$(2.16) \quad \bar{\Gamma}(\bar{x}, \bar{\xi}_1(\bar{x}), \bar{\xi}_2(\bar{x})) = \bar{x}(x; \Gamma(x, \xi_1(x), \xi_2(x))) - \bar{x}(x; \xi_1(x); \xi_2(x)).$$

We shall use these laws of transformation in the next section.

3. Covariant differential of multilinear forms in covariant and contravariant vector fields. Since the covariant differential of a covariant vector is a covariant vector depending linearly on an arbitrary contravariant vector, it is clear that the theory of successive covariant differentials can be brought under the theory of covariant differentials of multilinear forms. We shall prove

THEOREM 3.1. *If*

(i) *$F(x, \xi_1(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_s(x))$ is a covariant vector valued multilinear form in the continuously differentiable contravariant vectors $\xi_1(x), \cdots, \xi_r(x)$ and covariant vectors $\eta_1(x), \cdots, \eta_s(x)$,*

(ii) *the partial differential $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s; \delta x)$ exists continuous in x,*

then the function $F(x, \xi_1(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_s(x)/\delta x)$ defined by

$$(3.1) \quad \left\{ \begin{aligned} &F(x, \xi_1(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_s(x)/\delta x) \\ &\quad = F(x, \xi_1(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_s(x); \delta x) \\ &\qquad - \sum_{i=1}^{r} F(x, \xi_1(x), \cdots, \xi_{i-1}(x), \Gamma(x, \xi_i(x), \delta x), \\ &\qquad\qquad\qquad \xi_{i+1}(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_s(x)) \\ &\qquad + \sum_{i=1}^{s} F(x, \xi_1(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_{i-1}(x), \\ &\qquad\qquad\qquad L(x, \eta_i(x), \delta x), \eta_{i+1}(x), \cdots, \eta_s(x)) \\ &\qquad - L(x, F(x, \xi_1(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_s(x)), \delta x) \end{aligned} \right.$$

---

[9] Michal (1), (2).

[10] Michal (1).

is a covariant vector valued multilinear form in $\xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s, \delta x$. *We shall call* $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s/\delta x)$ *the covariant differential of* $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s)$.

Proof. We shall give the details of proof for $r = 1$, $s = 1$ as the proof for the general case, although lengthy, differs in no essential manner from that of this special case.

By hypothesis

$$\bar{F}\big(\bar{x}, \bar{\xi}(\bar{x}), \bar{\eta}(\bar{x})\big) = x^*\big(\bar{x}; F(x, \xi(x), \eta(x))\big) \tag{3.2}$$

from which we obtain with the aid of Theorem 2.1

$$\left\{\begin{aligned} \bar{F}\big(\bar{x}, \bar{\xi}(\bar{x}), \bar{\eta}(\bar{x}); \delta\bar{x}\big) &= x^*\big(\bar{x}; F(x, \xi(x), \eta(x); \delta x)\big) \\ &\quad + x^*\big(\bar{x}; \delta\bar{x}; F(x, \xi(x), \eta(x))\big) \\ &\quad + x^*\big(\bar{x}; d_{\delta x}^{\overset{x}{\sigma}} F(x, \xi(\sigma), \eta(\sigma))\big) - d_{\delta\bar{x}}^{\overset{\bar{x}}{\sigma}} \bar{F}(\bar{x}, \bar{\xi}(\sigma), \bar{\eta}(\sigma)). \end{aligned}\right. \tag{3.3}$$

On using (2.16) and (3.2) we find that

$$\left\{\begin{aligned} \bar{F}\big(\bar{x}, \bar{\Gamma}(\bar{x}, \bar{\xi}(\bar{x}), \delta\bar{x}), \bar{\eta}(\bar{x})\big) &= x^*\big(\bar{x}; F(x, \Gamma(x, \xi(x), \delta x), \eta(x))\big) \\ &\quad - \bar{F}\big(\bar{x}, \bar{x}(x; \xi(x); \delta x), \bar{\eta}(\bar{x})\big). \end{aligned}\right. \tag{3.4}$$

Similarly from (2.10) and (3.2)

$$\left\{\begin{aligned} \bar{F}\big(\bar{x}, \bar{\xi}(\bar{x}), \bar{L}(\bar{x}, \bar{\eta}(\bar{x}), \delta\bar{x})\big) &= x^*\big(\bar{x}; F(x, \xi(x), L(x, \eta(x), \delta x))\big) \\ &\quad + \bar{F}\big(\bar{x}, \bar{\xi}(\bar{x}), x^*(\bar{x}; \delta\bar{x}; \eta(x))\big). \end{aligned}\right. \tag{3.5}$$

Evidently

$$\left\{\begin{aligned} \bar{L}\big(\bar{x}, \bar{F}(\bar{x}, \bar{\xi}(\bar{x}), \bar{\eta}(\bar{x})), \delta\bar{x}\big) &= x^*\big(\bar{x}; L(x, F(x, \xi(x), \eta(x)), \delta x)\big) \\ &\quad + x^*\big(\bar{x}; \delta\bar{x}; F(x, \xi(x), \eta(x))\big). \end{aligned}\right. \tag{3.6}$$

Taking the differential of (2.1) and using (3.2) we obtain

$$\left\{\begin{aligned} \bar{F}\big(\bar{x}, \delta\bar{\xi}(\bar{x}), \bar{\eta}(\bar{x})\big) &= \bar{F}\big(\bar{x}, \bar{x}(x; \xi(x); \delta x), \bar{\eta}(\bar{x})\big) \\ &\quad + x^*\big(\bar{x}; F(x, \delta\xi(x), \eta(x))\big). \end{aligned}\right. \tag{3.7}$$

Similarly from (2.11) and (3.2)

$$\left\{\begin{aligned} \bar{F}\big(\bar{x}, \bar{\xi}(\bar{x}), \delta\bar{\eta}(\bar{x})\big) &= \bar{F}\big(\bar{x}, \bar{\xi}(\bar{x}), x^*(\bar{x}; \delta\bar{x}; \eta(x))\big) \\ &\quad + x^*\big(\bar{x}; F(x, \xi(x), \delta\eta(x))\big). \end{aligned}\right. \tag{3.8}$$

Reducing (3.3) by means of (3.7) and (3.8)

$$(3.9)\quad \begin{cases} \bar{F}\big(\bar{x}, \bar{\xi}(\bar{x}), \bar{\eta}(\bar{x}); \delta\bar{x}\big) \\ \quad = x^*\big(\bar{x}; F(x, \xi(x), \eta(x); \delta x)\big) + x^*\big(\bar{x}; \delta\bar{x}; F(x, \xi(x), \eta(x))\big) \\ \qquad - \bar{F}\big(\bar{x}, \bar{x}(x; \xi(x); \delta x), \bar{\eta}(\bar{x})\big) - \bar{F}\big(\bar{x}, \bar{\xi}(\bar{x}), x^*(\bar{x}; \delta\bar{x}; \eta(x))\big). \end{cases}$$

Finally with the aid of (3.4), (3.5), (3.6), and (3.9) we obtain

$$\bar{F}\big(\bar{x}, \bar{\xi}(\bar{x}), \bar{\eta}(\bar{x})/\delta\bar{x}\big) = x^*\big(\bar{x}; F(x, \xi(x), \eta(x)/\delta x)\big),$$

which completes the proof of the special case $r = 1$, $s = 1$.

The following two theorems can now be proved without much difficulty.

THEOREM 3.2. *If in the hypotheses of Theorem 3.1, the function $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s)$ is taken to be a contravariant vector valued multilinear form, then the function $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s/\delta x)$ defined by*

$$(3.10)\quad \begin{cases} F\big(x, \xi_1(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_s(x)/\delta x\big) \\ \quad = F\big(x, \xi_1(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_s(x); \delta x\big) \\ \qquad - \sum_{i=1}^{r} F\big(x, \xi_1, \cdots, \xi_{i-1}, \Gamma(x, \xi_i(x), \delta x), \xi_{i+1}, \cdots, \xi_r, \eta_1, \cdots, \eta_s\big) \\ \qquad + \sum_{i=1}^{s} F\big(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_{i-1}, L(x, \eta_i, \delta x), \eta_{i+1}, \cdots, \eta_s\big) \\ \qquad + \Gamma\big(x, F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s), \delta x\big) \end{cases}$$

is a contravariant valued multilinear form in $\xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s, \delta x$. We shall call $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s/\delta x)$ the covariant diffierential of $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s)$.

THEOREM 3.3. *If in the hypotheses of Theorem 3.1, the function $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s)$ is taken to be an absolute scalar multilinear form (with numerical values or with values in a Banach space), then the function $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s/\delta x)$ defined by*

$$(3.11)\quad \begin{cases} F\big(x, \xi_1(x), \cdots, \xi_r(x), \eta_1(x), \cdots, \eta_s(x)/\delta x\big) \\ \quad = F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s; \delta x) \\ \qquad - \sum_{i=1}^{r} F\big(x, \xi_1, \cdots, \xi_{i-1}, \Gamma(x, \xi_i, \delta x), \xi_{i+1}, \cdots, \xi_r, \eta_1, \cdots, \eta_s\big) \\ \qquad + \sum_{i=1}^{s} F\big(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_{i-1}, L(x, \eta_i, \delta x), \eta_{i+1}, \cdots, \eta_s\big) \end{cases}$$

is an absolute scalar multilinear form (with numerical or Banach values) in $\xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s, \delta x$. We shall call $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s/\delta x)$ the covariant differential of $F(x, \xi_1, \cdots, \xi_r, \eta_1, \cdots, \eta_s)$.

To have the successive covariant differential $\eta(x/\delta_1 x/\cdots/\delta_p x)$ well defined in all coördinate systems, it is sufficient to assume that (1) the covariant vector $\eta(x)$ possesses a continuous p-th differential; (2) $L(x,\eta,\delta x)$ has a continuous p-th partial differential in the first place; (3) $\Gamma(x,\xi_1,\xi_2)$ has a continuous $(p-1)$-st partial differential in the first place; (4) $\bar{x}(x)$ has a continuous $(p+2)$-nd differential; and (5) $x^*(\bar{x};\lambda)$ has a continuous $(p+1)$-st differential in $\bar{x}$.

By calculation we obtain the commutation rule

$$(3.12)\quad \begin{aligned}&\eta(x/\delta_1 x/\delta_2 x)-\eta(x/\delta_2 x/\delta_1 x)\\ &\qquad=-L(x,\eta(x),\delta_1 x,\delta_2 x)-2\eta(x/\Omega(x,\delta_1 x,\delta_2 x)),\end{aligned}$$

where

$$(3.13)\quad \left\{\begin{aligned}L(x,\eta(x),\delta_1 x,\delta_2 x)&=L(x,\eta(x),\delta_1 x;\delta_2 x)-L(x,\eta(x),\delta_2 x;\delta_1 x)\\ &\quad+L\big(x,L(x,\eta(x),\delta_2 x),\delta_1 x\big)-L\big(x,L(x,\eta(x),\delta_1 x),\delta_2 x\big)\end{aligned}\right.$$

and

$$(3.14)\qquad \Omega(x,\delta_1 x,\delta_2 x)=\tfrac{1}{2}\{\Gamma(x,\delta_1 x,\delta_2 x)-\Gamma(x,\delta_2 x,\delta_1 x)\}.$$

Since $\Omega(x,\delta_1 x,\delta_2 x)$ is the contravariant vector valued *torsion* form, it follows from (3.12) and Theorem 3.1 that the trilinear form $L(x,\eta(x),\delta_1 x,\delta_2 x)$ is a covariant vector valued trilinear form, called the (covariant vector valued) *curvature form.*

Suppose now that $F(x,\eta(x),\delta x)$ is bilinear in covariant vectors $\eta(x)$ and in δx, and suppose further that $F(x,\eta(x),\delta x)$ is the component of a geometric object. On making special use of the properties of the adjoints and the law of transformation of the linear connection one can demonstrate without much difficulty the following theorem.

THEOREM 3.4. *A necessary and sufficient condition that the adjointness relation*

$$[\Gamma(x,\xi(x),\delta x),\eta(x)]=[\xi(x),F(x,\eta(x),\delta x)]$$

be a geometric condition (i.e., continues to hold under a transformation of coördinates) is that $F(x,\eta(x),\delta x)$ be a covariant linear connection.

The importance of a relation of the above type between the linear connection Γ and the covariant linear connection L is made clear from the following result.

THEOREM 3.5. *A necessary and sufficient condition that*

$$\delta[\xi(x),\eta(x)]=[\xi(x/\delta x),\eta(x)]+[\xi(x),\eta(x/\delta x)]$$

for all continuously differentiable contravariant vectors $\xi(x)$ *and covariant vectors* $\eta(x)$ *is that the covariant linear connection* $L(x, \eta(x), \delta x)$ *be the adjoint of the linear connection* $\Gamma(x, \xi(x), \delta x)$ *considered as a linear function of* $\xi(x)$.

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIFORNIA.

REFERENCES.

A. D. Michal, I. E. Highberg and A. E. Taylor:
(1) "Abstract Euclidean spaces with independently postulated analytical and geometrical metrics," *Annali di Pisa* (in press).

A. D. Michal:
(1) "General tensor analysis," *Bulletin of the American Mathematical Society* (in press).
(2) "Postulates for a linear connection," *Annali di Matematica* (in press).

A. D. Michal and V. Elconin:
(1) "Completely integrable differential equations in abstract spaces," *Acta Mathematica* (in press).

A TYPE OF HOMOGENEITY FOR CONTINUOUS CURVES.[1]

By Charles H. Wheeler, III.

Introduction. A set of points M is said to be *homogeneous*[2] if, given any two points x and y of M, it is possible to find a homeomorphism which will send M into itself in such a way that x is sent into y.

A set of points M is said to be *bi-homogeneous*[3] if given any two points x and y of M there exists a homeomorphism which sends M into itself in such a way that x is sent into y and y is sent into x.

We will investigate the conditions under which a compact, locally connected continuum may be *cyclic element homogeneous,* i. e., given any two true cyclic elements[4] of a set M, there exists a homeomorphism which sends M into itself in such a way that one of the given true cyclic elements is sent into the other. A compact locally connected continuum is a continuous curve in the sense that it is a set of points which is the image of the unit interval under a continuous transformation.

The case where M contains only a finite number of true cyclic elements will be completely treated, while some results will be stated for the case where M contains infinitely many true cyclic elements.

Two simple closed curves joined by a simple arc provide an example of a cyclic element homogeneous set. The curve illustrated in Fig. 1 is not cyclic element homogeneous because the true cyclic element marked C_1 cannot be sent into any of the other true cyclic elements by a homeomorphism which sends the set into itself. In Fig. 2, also in Fig. 1, each true cyclic element is homeomorphic with each of the remaining ones, but in Fig. 2 the cyclic element marked C_2 cannot be sent into any of the others by a homeomorphism which sends the set into itself. Fig. 3 is an example of a set of points which contains an infinite number of true cyclic elements and is cyclic element homogeneous.

---

[1] Received April 18, 1935; revised October 15, 1936.

[2] See Kuratowski, *Fundamenta Mathematicae,* T. 3 (1922), pp. 14-19, also Mazurkiewicz, *Fundamenta Mathematicae,* T. 5 (1924), pp. 137-146.

[3] Kuratowski, *loc. cit.*

[4] The cyclic elements of a locally connected continuum are (1) all cut points of the continuum and (2) the set of all points conjugate to a point p, where p is any non-cut point of the continuum. A true cyclic element is one which does not reduce to a single point. See G. T. Whyburn, "Concerning the structure of a continuous curve," *American Journal of Mathematics,* vol. 50 (1928), pp. 167-194, and Kuratowski and Whyburn, "Sur les eléménts cycliques et leurs applications," *Fundamenta Mathematicae,* T. 16 (1930), pp. 305-331.

1. *Preliminary theorems.* Let M be any compact and locally connected continuum, which we shall consider as a space. Designate by H the smallest A-set[5] which contains all the true cyclic elements C_i in M, i. e., an A-set containing all the true cyclic elements C_i and not a proper subset of any A-set

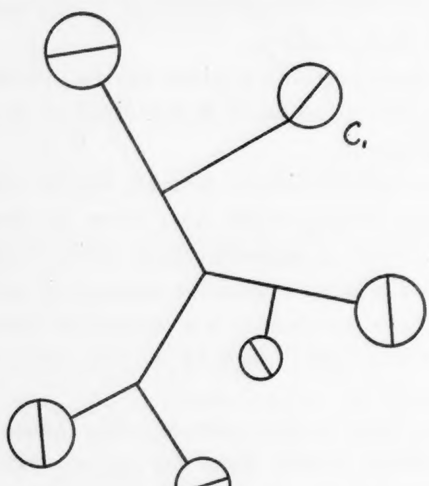


Fig. 1

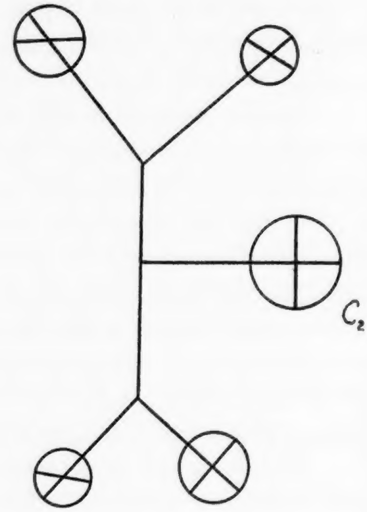


Fig. 2

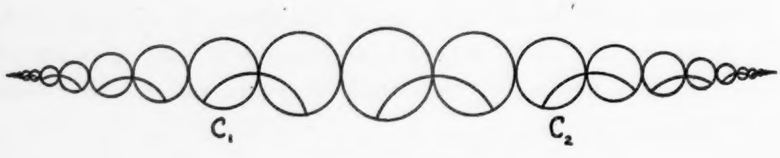


Fig. 3

containing all the true cyclic elements. The set H may be obtained by taking the product of all A-sets in M which contain all the true cyclic elements. Then since[6] the product of any number of A-sets is an A-set, it follows H is an

[5] A closed set which has the property that if $x, y \in A$ then every arc xy in the space is contained in A.

[6] Cf. Kuratowski and Whyburn, *loc. cit.*, Theorem 4: 1.

A-set and clearly it is the smallest such set containing all the true cyclic elements. If there are only a finite number of true cyclic elements, the set H may be obtained by choosing a non-cut point p_i from each true cyclic element C_i, and forming all the possible pairs of these points. Then for each pair p_i, p_j form the cyclic chain $C_k(p_i, p_j)$. We then have $H = \sum_1^n C_k(p_i, p_j)$.

1.1. THEOREM. *If M contains only a finite number (> 1) of true cyclic elements, then there exists at least two true cyclic elements each of which has only one point in common with the closure of the remainder of H.*

Proof. This follows immediately from a theorem of G. T. Whyburn.[7] He proves that if a continuum has more than one cyclic element then it contains at least two nodes, where a node is defined as an end point or a true cyclic element which contains only one cut point. Hence H must contain at least two true cyclic elements which contain only one cut point since it contains no end points.

1.2. THEOREM. *If T is any homeomorphism which sends M into itself, then $T(H) = H$.*

Proof. The set H is uniquely defined as the smallest A-set containing all the true cyclic elements of M. Any homeomorphism T will send H into an A-set which contains all the true cyclic elements of M, thus $H \subset T(H)$.

Since T^{-1} is a homeomorphism, $H \subset T^{-1}(H)$. Operating upon this with T we have $T(H) \subset TT^{-1}(H) = H$. Thus $T(H) \subset H$. It follows from the above two inclusions that $T(H) = H$, and the theorem is proved.

We will consider H to be our space for the remainder of this paper. H is a compact locally connected continuum.

1.3. *Definition.* A set H is said to be *cyclic element homogeneous* if, given any two true cyclic elements of H, then there exists a homeomorphism T such that $T(H) = H$ and one of the given true cyclic elements is sent into the other.

From this definition we have immediately the following:

1.4. THEOREM. *If H is cyclic element homogeneous, then each true cyclic element contains the same number (finite or infinite) of cut points of H.*

2. *The case where H contains only a finite number of true cyclic elements.* If the set H contains only a finite number (> 1) of true cyclic elements,

[7] "Concerning the structure of a continuous curve," *loc. cit.*, p. 180.

there exists a true cyclic element C_s which contains only one point which cuts H by 1. 1. Then every true cyclic element of H has only one point in common with the closure of the remainder of H, by 1. 3, i. e., $C_i \cdot \overline{H - C_i}$ = a single point.

Definition. Let $p_1{}^i$ be the point of C_i such that $p_1{}^i \in \overline{H - C_i}$. The sum of all true cyclic elements which contain $p_1{}^i$ is called a *cluster* of true cyclic elements. The point $p_1{}^i = p_j$ is called the center of the cluster K_j.

Then $K_j = \sum_{i=1}^{n} C_i{}^j$ and $\prod_{i=1}^{n} C_i{}^j = p_j$. The centers of every two clusters are joined by a simple arc in H, since H is an A-set. The arc p_jp_k is such that

$$p_jp_k \cdot K_j = p_j \quad \text{and} \quad p_jp_k \cdot K_k = p_k,$$

for we have seen above that no $C_i{}^j$ of K_j can have more than one point in common with $\overline{H - C_i{}^j}$.

2. 1. Theorem. *If H is cyclic element homogeneous and contains only a finite number of true cyclic elements, then there is the same number of true cyclic elements in each cluster.*

Proof. Suppose the contrary, then there exists some two clusters such that

$$K_1 = \sum_1^n C_i{}^1, \qquad K_2 = \sum_1^m C_i{}^2, \qquad n > m.$$

The points p_1 and p_2 are the centers of the clusters K_1 and K_2 respectively. Since H is cyclic element homogeneous, let $T(H) = H$ in such a way that $T(C_1{}^1) = C_1{}^2$. Now $C_1{}^1 \cdot \overline{K_1 - C_1{}^1} = p_1$ and $C_1{}^2 \cdot \overline{K_2 - C_1{}^2} = p_2$, and hence $T(p_1) = p_2$.

Since $C_2{}^1 \supset p_1$, then $T(C_2{}^1)$ = some $C_i{}^2$ for some i, $i = 2, 3, \cdots, m$. This is also true for the remaining $n - 2$ true cyclic elements in K_1, but there are $m - n$ less true cyclic elements in K_2 than in K_1, so this is impossible. Thus the supposition that the clusters do not contain the same number of true cyclic elements leads to a contradiction, and the theorem is proved.

2. 2. Corollary. *If a true cyclic element of one cluster is transformed by a homeomorphism T into a true cyclic element of another cluster, then the first cluster is transformed into the second cluster by T.*

2. 3. Theorem. *If H is cyclic element homogeneous and contains only a finite number of true cyclic elements, then no cluster cuts H.*

Proof. Suppose there exists no cluster which does not cut H. There are only a finite number of clusters since there are only a finite number of true cyclic elements in H. Then $H - K_1 = H_1^1 + H_2^1$, where

$$H_1^1 \cdot \overline{H_2^1} = \overline{H_1^1} \cdot H_2^1 = 0 \quad \text{and} \quad H_1^1 \neq 0 \neq H_2^1.$$

H_1^1 contains at least one cluster K_{i_1}. Let K_{n_2} be the cluster in H_1^1 with the least subscript. Then $H - K_{n_2} = H_1^2 + H_2^2$, where

$$H_1^2 \cdot \overline{H_2^2} = \overline{H_1^2} \cdot H_2^2 = 0, \quad H_1^2 \neq 0 \neq H_2^2 \quad \text{and} \quad H_2^2 \supset K_1.$$

H_1^2 contains at least one cluster K_{i_2}, $K_{i_2} \neq K_{n_2}$. Let K_{n_3} be the cluster in H_1^2 with the least subscript. Then $H - K_{n_3} = H_1^3 + H_2^3$, where

$$H_1^3 \cdot \overline{H_2^3} = \overline{H_1^3} \cdot H_2^3 = 0, \quad H_1^3 \neq 0 \neq H_2^3 \quad \text{and} \quad H_2^3 \supset K_{n_2}.$$

H_1^3 contains at least one cluster K_{i_3}, $K_{i_3} \neq K_{n_3}$.

This can be carried on indefinitely, but this is impossible for there are only a finite number of clusters. Thus there exists at least one cluster K_t which does not cut H.

Let K_p be any cluster of H different from K_t. By 2.2 $H - K_t$ is homeomorphic with $H - K_p$ and hence K_p can not cut H. Thus no cluster cuts H.

2.4. *Definitions.*

2.41. A compact locally connected continuum H is said to be *symmetrical* with respect to a cyclic element C (whether C be a true cyclic element or not) if every component of $H - C$ is homeomorphic with every other component of $H - C$. We will call C the center of symmetry.

2.42. A set of points H is said to be *cyclic symmetrical* with respect to a cyclic element C (whether C be a true cyclic element or not) if it is symmetrical and any true cyclic element of one component of $H - C$ may be sent into any true cyclic element of another component of $H - C$ by a homeomorphism which sends the first component into the second.

2.43. A *major branch* at a branch point x is the component of $H - x$ which contains the center of symmetry.

2.44. A *minor branch* at a branch point x is a component of $H - x$ which does not contain the center of symmetry.

2.5. THEOREM. *If H is cyclic element homogeneous and contains only a finite number of true cyclic elements, then there exists a point c such that H is cyclic symmetrical with respect to c.*

Proof. We saw that the centers of every pair of clusters are joined by an arc in H, and that this arc has only the centers of the clusters in common with the clusters. Also by 2.3 no cluster cuts H. Thus $\overline{H - \sum_1^m K_i}$ is an acyclic curve with a finite number of branches.

There is a center p_i of a cluster K_i at the end of each branch. Form all the possible pairs of the points p_i. Take a pair p_1, p_2 which has a maximum number of branch points on the arc joining them. If this number is odd let the middle branch point be c, if it is even let c be any point on the open arc joining the two middle branch points. The point c is uniquely determined in the case where the maximum number is odd, and uniquely determined to the extent of being any one of the points on an open arc in the case where the maximum number is even. Now

$$H - c = S_1 + S_2 + S_3 + \cdots + S_n,$$

where $S_1, S_2, \cdots, S_n$ are components. If the number of branch points on the arc joining p_1 and p_2 is even there are only two components, while if this number is odd there may be any finite number of components. Let $S_1 \supset p_1$ and $S_2 \supset p_2$.

Number the branch points on S_1 in the following way: Let the first branch point out from c be x^1, then on each minor branch from x^1 let the first branch points be $x_1^2, x_2^2, \cdots, x_m^2$. The branch point on the branch containing p_1 we will call x^2. Let the first branch points on the minor branches from x_1^2 be $x^3_{1,1}, x^3_{1,2}, \cdots, x^3_{1,k}$. The branch point on the branch containing p_1 we will call x^3. Continue in this manner until all the branch points have been numbered. Number the branch points on S_2 in the same manner except with the use of y instead of x.

The number of branch points from c to p_1 is the same as the number from c to p_2; let us assume this number is k.

We must now show that there exists a homeomorphism T such that $T(S_1) = S_2$. Since H is cyclic element homogeneous, let T be a homeomorphism of H into itself such that $T(C_1^1) = C_1^2$ where $C_1^1 \supset p_1$ and $C_1^2 \supset p_2$. Then $T(p_1) = p_2$ and $T(K_1) = K_2$. The arc p_1x^k is sent into the arc p_2y^k and x^k into y^k by T. At the branch points x^k and y^k there are the same number of branches because x^k and y^k correspond to each other under a homeomorphism. There is no branch point on any of the minor branches at x^k or y^k for if there were one on some minor branch, say at y^k, there would be more branch points on the arc joining one of the centers p_i (on this branch) and p_1 than on the arc p_1p_2. Then p_1, p_2 would not be a maximal pair.

The arc $x^k x^{k-1}$ is sent into the arc $y^k y^{k-1}$ by T so that x^{k-1} is sent into y^{k-1}. There are the same number of branches at x^{k-1} as at y^{k-1} because x^{k-1} and y^{k-1} corrrespond to each other under a homeomorphism. There can not be more than one branch point on any of the minor branches at x^{k-1} or y^{k-1}; for if there were, p_1, p_2 would not have been a maximal pair. There must be one branch point on each of the minor branches at x^{k-1} and y^{k-1} and the number of minor branches at these branch points is the same as the number of minor branches at x^k or y^k.

Now continue down the major branch at x^{k-1} and y^{k-1}. The arc $x^{k-1}x^{k-2}$ is sent into the arc $y^{k-1}y^{k-2}$ by T, and then x^{k-2} is sent into y^{k-2}. The number of minor branches at x^{k-2} and y^{k-2} is the same, because x^{k-2} and y^{k-2} correspond to each other under a homeomorphism. The minor branches at x^{k-2} are homeomorphic with the minor branches at y^{k-2}. On continuing in this manner until we reach $T(x^1) = y^1$, it can be shown that the minor branches at x^1 are homeomorphic with the minor branches at y^1. Then the arc $x^1c - c$ is sent into the arc $y^1c - c$ by T. Therefore $T(S_1) = S_2$.

If there are only two components of $H - c$, the theorem is proved. In the case where there are more than two components of $H - c$, take any component S_3 different from S_1 and S_2. Let p_3 be the center of the cluster K_3 such that the number of branch points on the arc joining it to c is a maximum for all centers p_i in S_3. Let the number of branch points on this arc be m. Then $m \leq k$, for if it were greater p_1, p_2 would not have been a maximal pair. Number the branch points on S_3 the same as on S_1, denoting them by z instead of x.

Since H is cyclic element homogeneous, let $T(H) = H$ in such a way that $T(C_1^1) = C_1^3$, where $C_1^1 \subset K_1 \subset S_1$ and $C_1^3 \subset K_3 \subset S_3$. Then $T(K_1) = K_3$, $T(p_1) = p_3$, $T(p_1x^k) = p_3z^m$ and $T(x^k) = z^m$. The number of branches at x^k is the same as the number at z^m, for x^k and z^m correspond to each other under a homeomorphism. There is no branch point on any of the minor branches at x^k or z^m, for if there were, p_3 would not have been a maximal number of branch points from c. Then $T(x^k x^{k-1}) = z^m z^{m-1}$ and $T(x^{k-1}) = z^{m-1}$. The number of branch points at x^{k-1} and z^{m-1} are the same since x^{k-1} and z^{m-1} correspond under a homeomorphism. The minor branches at z^{m-1} are homeomorphic and homeomorphic to the minor branches at x^{k-1}.

It is thus seen that if $m = k$, S_1 is homeomorphic with S_3. If $m < k$, say $m + 1 = k$, then one minor branch at x^1 is sent into S_3 and x^1 is sent into c. Since x^1 and c correspond to each other under a homeomorphism, there are the same number of branches at x^1 as there are at c. This may or may not be true; if not, it is seen immediately that $T(H) \neq H$, and therefore

$m = k$. If the number of branches at x^1 and c are the same, consider the number of branch points on the arc from x^1 to $p_2 \epsilon S_2$. There are $k+1$. The branch from x^1 containing these must be sent into a component of $H - c$ different from S_3; but we have seen that the maximum number of branch points from c to the center of any cluster in $H - c$ was k. Therefore this is impossible and $T(H) \neq H$. Thus $m = k$. Therefore all of the components of $H - c$ are homeomorphic.

It must now be shown that any true cyclic element in one component may be sent into any true cyclic element in another component by a homeomorphism which sends the first component into the second. Let $C_1{}^i$ be any true cyclic element in S_r and $C_1{}^j$ any in S_t. Now $C_1{}^i \subset K_i$ and $C_1{}^j \subset K_j$. Since H is cyclic element homogeneous, let $T(H) = H$ in such a manner that $T(C_1{}^i) = C_1{}^j$. It has already been shown that S_r and S_t are homeomorphic, that every cluster on S_r and S_t are the same number (k) of branch points away from c, that at each i-th branch point out from c ($i = 1, 2, 3, \cdots, k$) there are the same number of minor branches which are homeomorphic with one another. Let the branch points on S_r and S_t be numbered in the same manner as were the branch points on S_1 and denoted by u and v, respectively. We have then $T(K_i) = K_j$, $T(p_i) = p_j$, $T(p_i u^k) = p_j v^k$ and $T(u^k) = v^k$. The remaining minor branches at u^k are sent into the remaining minor branches at v^k by T. Then $T(u^k u^{k-1}) = v^k v^{k-1}$ and $T(u^{k-1}) = v^{k-1}$. The remaining minor branches at u^{k-1} are sent into the remaining minor branches at v^{k-1} by T. Continuing this finally yields $T(u^1) = v^1$. The remaining minor branches at u^1 are sent into the remaining minor branches at v^1. Then $T(u^1 c - c) = v^1 c - c$, thus $T(S_r) = S_t$. Q. E. D.

2. 6. *Summary.* We have proved that if H is cyclic element homogeneous and contains only a finite number of true cyclic elements C_i then

1) $C_i \cdot \overline{H - C_i} =$ a single point, for each i.
2) Each cluster K_i has the same number of true cyclic elements.
3) $H - K_i$, for every i, is connected.
4) $\overline{H - \sum_1^n K_i}$ is an acyclic curve with a finite number of branches.
5) There exists a point c such that
   a. Every cluster is the same number of branch points distant from c.
   b. At each i-th branch point from c, ($i = 1, 2, \cdots, k$) there are the same number of branches, and all the minor branches are homeomorphic.
   c. On every component of $H - c$ there are the same number of branch points and the same number of clusters.

d. Every component of $H - c$ is homeomorphic with every other and any true cyclic element of one component may be sent into any true cyclic element of any other component by a homeomorphism T which sends the first component into the second and which sends H into itself.

6) All the true cyclic elements are homeomorphic, and if $p_i^r \epsilon C_r$ and $p_i^s \epsilon C_s$ such that $C_r \cdot \overline{H - C_r} \supset p_i^r$ and $C_s \cdot \overline{H - C_s} \supset p_i^s$, then the homeomorphism $W(C_r) = C_s$ is such that $W(p_i^r) = p_i^s$ for some i.

Fig. 4 is an example of a set L which is cyclic element homogeneous. The point c is the center of symmetry, and there are three components of $L - c$.

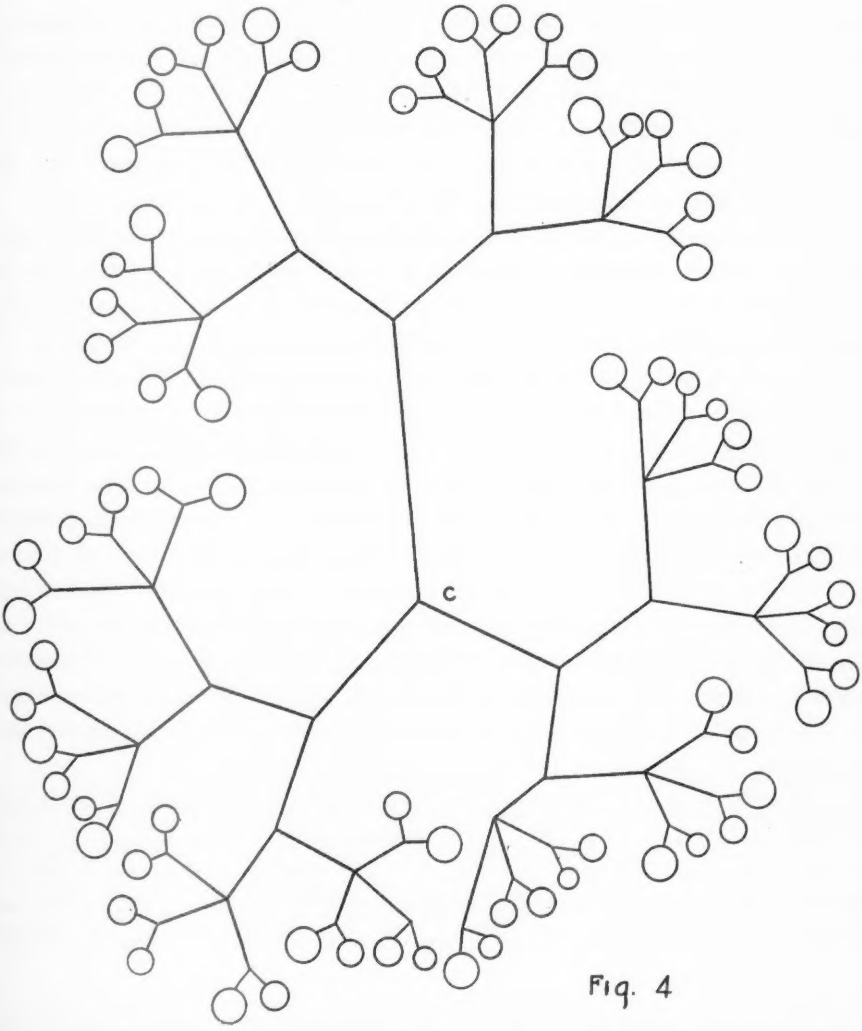


Fig. 4

2. 7. THEOREM. *If H contains only a finite number of true cyclic elements, in order that H be cyclic element homogeneous it is necessary and sufficient that* (1) *the true cyclic elements be grouped in clusters with the same number in each cluster,* (2) *no cluster cuts H,* (3) *there exists a point c of H such that each cluster is the same number, k, of branch points away from c,* (4) *at each i-th branch point* $(i = 1, 2, \cdots, k)$ *from c there are the same number of branches, and* (5) *if C_r and C_s are any two true cyclic elements of H, there exists a homeomorphism of C_r into C_s which sends the cut points of H on C_r into the cut points of H on C_s.*

Proof. The necessity follows from 2. 5. To show the sufficiency take any two components S_1, S_2 of $H - c$. Denote the branch points on S_1 by x and those on S_2 by y, as was done in 2. 5. Go out along the arcs from c to x^1 and to y^1 lying in S_1 and S_2 respectively. There are the same number of branches at x^1 as at y^1, by hypothesis. Out from x^1 and y^1 on each minor branch there is a second branch point from c. There are the same number of branches at each of these by hypothesis. Out from the second branch points from c on each of the minor branches is the third branch point from c, and there are the same number of branches at each of these.

Continue this until we get to the k-th branch points from c. By hypothesis there are the same number of branches at each of these points. Out from the k-th branch points on each of the minor branches is a cluster of true cyclic elements, for otherwise this branch would not have been in H. There can not be any branch point on any of these branches, for if there were there would be at least two clusters which had more than k branch points on the arc from the center of the cluster to c. There is no cluster on any minor branch from x^1 and y^1 before the k-th branch point, for if there were there would be less than k branch points on the arc from its center to c. Thus there is a cluster of true cyclic elements at the end of each minor branch from the k-th branch points. By hypothesis there is the same number of true cyclic elements in each cluster. Since no cluster cuts H, every two of the clusters are joined by an arc in H which passes through at least one branch point. Each true cyclic element contains only one point which cuts H, namely the center of the cluster in which the true cyclic element is contained. This follows from the fact that no cluster cuts H.

It must now be shown that there exists a homeomorphism T such that if there be given any two true cyclic elements C_i, C_j of H, then $T(H) = H$ in such a way that $T(C_i) = C_j$. Take any two true cyclic elements C_1 and C_2. By hypothesis there exists a homeomorphism W such that $W(C_1) = C_2$ and $W(p_1) = p_2$ where $p_1 \epsilon C_1 \cdot \overline{H - C_1}$ and $p_2 \epsilon C_2 \cdot \overline{H - C_2}$. Define the homeo-

morphism $T = W$ over C_1. There are the same number of true cyclic elements in each cluster, hence the definition of T can be extended so that $T(K_1) = K_2$ where $K_1 \supset C_1$ and $K_2 \supset C_2$. If p_1x^k and p_2y^k are the arcs from K_1 and K_2 to the branch points x^k and y^k respectively, we so extend the definition of T that $T(p_1x^k) = p_2y^k$, and $T(x^k) = y^k$; also $T(b^i{}_{x,k}) = b^i{}_{y,k}$ where $b^i{}_{x,k}$ are the minor branches at x^k which do not contain p_1 and $b^i{}_{y,k}$ the minor branches at y^k which do not contain p_2, $i = 2, 3, \cdots, m$. If x^kx^{k-1} and y^ky^{k-1} are the arcs from x^k and y^k to the $(k-1)$-st branch points, we define $T(x^kx^{k-1}) = y^ky^{k-1}$ and $T(x^{k-1}) = y^{k-1}$; also $T(b^i{}_{x,k-1}) = b^i{}_{y,k-1}$ where $b^i{}_{x,k-1}$ are the minor branches at x^{k-1} not containing x^k and $b^i{}_{y,k-1}$ the minor branches at y^{k-1} not containing y^k, $i = 2, 3, \cdots, j$. Continue in this manner until $x^j = y^j$ or until c is reached on both components. If b^1 and b^2 are the branches from x^j or c, as the case may be, which contain C_1 and C_2 respectively, for each point $x \in b^2$, define $T(x) = T^{-1}(x)$; for each point $x \in H - b^1 - b^2$, define $T(x) = x$. We thus have a homeomorphism T which sends C_1 into C_2, the component of $H - c$ which contains C_1 into the component of $H - c$ which contains C_2, and H into itself.

2. 8. *Definition.* A set of points H is said to be *bi-cyclic element homogeneous* if, given any two true cyclic elements of H, there exists a homeomorphism T such that $T(H) = H$ and the true cyclic elements are sent into each other.

From the way the homeomorphism T was defined in 2. 7, it is seen that if a set H satisfies the conditions of 2. 7 it is bi-cyclic element homogeneous.

As was stated in the introduction, Fig. 3 is an example of a space M which is cyclic element homogeneous but it is not bi-cyclic element homogeneous. The true cyclic element C_1 may be sent into any other true cyclic element of the space by a homeomorphism which sends M into itself; but C_1 can not be sent into C_2 by a homeomorphism which sends C_2 into C_1 and M into itself.

It may be remarked that the finite case just treated could have been reduced to the consideration of an acyclic curve which was end point homogeneous. However, little if any advantage in simplicity seems to accrue from such a reduction.

3. *The case of infinitely many true cyclic elements.* Although no complete solution for this case of the problem has yet been obtained, we shall state here some results bearing on certain important phases of it.

3. 1. If H is cyclic element homogeneous and contains infinitely many true cyclic elements which are grouped in clusters, where no cluster cuts H,

then there are a finite number of clusters with an infinite number of true cyclic elements in each cluster. Also under this hypothesis there exists a point c such that each cluster is the same number of branch points away from c and at each i-th branch point $(i = 1, 2, \cdots, k)$ from c there are the same number of branches. It can be shown that the necessary and sufficient condition for this case is similar to that of the finite case.

3.2. Let H satisfy the conditions: (1) it contains infinitely many true cyclic elements $\{C_j\}$, (2) each cut point is contained in exactly two true cyclic elements, (3) each component of $H - C_j$ has a different boundary point, (4) the set of cut points of H is totally disconnected, and (5) no point p is the limit of true cyclic elements in more than one component of $H - p$.

Under these restrictions, a necessary condition that H be cyclic element homogeneous is that when C_s and C_t are any two true cyclic elements of H, there exists a homeomorphism of C_s into C_t such that one particular cut point in C_s is sent into one particular cut point in C_t and the remaining cut points in C_s are sent into the remaining cut points in C_t.

Thus far it has not been shown that this condition is sufficient for H to be cyclic element homogeneous. But a sufficient condition is that when C_s and C_t are any two true cyclic elements of H, there exists a homeomorphism T of C_s into C_t such that two particular cut points in C_s are sent into two particular cut points in C_t and the remaining cut points in C_s are sent into the remaining cut points in C_t.

It is to be noted that the sufficient condition is stronger than the necessary condition.

THE JOHNS HOPKINS UNIVERSITY.

A NAVIGATION PROBLEM IN THE CALCULUS OF VARIATIONS.*

By E. J. McShane.

The problem which I shall consider in this note is closely related to the Zermelo Navigation problem.[1] Let us suppose that the velocities relative to the air which can be attained by an airship consist of all the vectors r lying in a convex body $K(x,t)$, depending on the position $x = (x^1, x^2, x^3)$ and the time t. The air is supposed to be in motion, its velocity being a continuous vector function $u(x,t)$. Given two points x_0, x_1, the problem is to find a path from x_0 to x_1 which can be traversed by the ship in the least possible time. If $K(x,t)$ is the sphere[2] $|r| \leqq k$, where k is a constant, and if we add the further requirement that the speed relative to the air shall almost always be exactly k, this becomes the Zermelo navigation problem. Our replacement of the sphere $|r| \leqq k$ by the convex body K is suggested by the fact that an airplane can travel faster down than up.

In the present paper I first prove under weak hypotheses the existence of a solution of the problem proposed above. I then consider the problem modified so as to be a generalization of the Zermelo problem (i. e. the ship's velocity is required to be almost always as great as possible), and under stronger hypotheses I prove that this problem also is solvable.

1. Throughout the following pages we shall use the following definitions and assumptions:

A is a bounded closed point set (atmosphere) in (x^1, x^2, x^3)-space.

Δ is a bounded closed set of real numbers t.

$K(x,t)$ is a bounded closed convex point set (or set of vectors) in three-dimensional space, defined and continuous[3] for $x \varepsilon A$ and $-\infty < t < \infty$.

$u(x,t)$ is a vector function $(u^1(x,t), u^2(x,t), u^3(x,t))$ defined and continuous for $x \varepsilon A$ and all t.

* Received December 3, 1936.

[1] For a detailed study of this problem, as well as a bibliography of previous papers, the reader is referred to a memoir by B. Manià, shortly to appear in *Mathematische Annalen*.

[2] $|r|$ is the length $[\Sigma(r^i)^2]^{1/2}$ of the vector r.

[3] For any convex set K, let K_ϵ be the set of all points having distance $\leqq \epsilon$ from K. A convex set $K(\theta)$ is a continuous function of θ at θ_0 if there is a neighborhood U of θ_0 such that for every θ in U the inclusions $K(\theta_0) < K_\epsilon(\theta)$ and $K(\theta) < K_\epsilon(\theta_0)$ hold.

$V(x, t)$ is the set of all vectors of the form $u(x, t) + r$ with $r \varepsilon K(x, t)$. That is, V is the translation of K by the vector u, and so it satisfies all the conditions imposed above on K.

Let the path of the ship be given in the form $C: x = x(t)$, $a \leqq t \leqq a + T$, where the parameter t is the time. The mean velocity of the ship between times t_1 and t_2 is $(x(t_2) - x(t_1))/(t_2 - t_1)$, and we shall assume that this is bounded. The velocity at time t is $x'(t)$, if $x'(t)$ exists. The velocity relative to the air is then $x'(t) - u(x(t), t)$, and this must be with the class $K(x(t), t)$ of velocities attainable at place $x(t)$ and time t. That is, by the definition of $V(x, t)$ we must have $x'(t) \varepsilon V(x(t), t)$. Combining this with the previous requirements on the paths $x(t)$ to be considered, we are led to the definition:

The curve $C: x = x(t)$, $a \leqq t \leqq a + T$ is *admissible* (or, more fully, an admissible curve traversable in the interval $[a, a + T]$) if

(1. 1a) the functions $x(t)$ are Lipschitzian;
(1. 1b) $x(t) \varepsilon A$ for $a \leqq t \leqq a + T$;
(1. 1c) $a \varepsilon \Delta$;
(1. 1d) $x(a) = x_0$, $x(a + T) = x_1$;
(1. 1e) $x'(t) \varepsilon V(x(t), t)$ wherever $x'(t)$ is defined.

It is convenient for the proof to have also the notation of a *weakly admissible* curve. A curve C is *weakly admissible* if it satisfies (1. 1a, b, c, d) and satisfies (1. 1e) if we replace the words "wherever $x'(t)$ is defined" by "almost everywhere."

We obtain an obviously equivalent definition by changing parameters from t to $\tau = (t - a)/T$; a curve $C: x = x(\tau)$, $0 \leqq \tau \leqq 1$ is an admissible curve traversable in the interval $[a, a + T]$ if

(1. 2a) the functions $x(\tau)$ are Lipschitzian;
(1. 2b) $x(\tau) \varepsilon A$ for $0 \leqq \tau \leqq 1$;
(1. 2c) $a \varepsilon \Delta$;
(1. 2d) $x(0) = x_0$, $x(1) = x_1$;
(1. 2e) $x'(\tau)/T \varepsilon V(x(\tau), a + T\tau)$ wherever $x'(\tau)$ is defined.

Likewise, if in (1. 2) we replace the words "wherever $x'(\tau)$ is defined" by "for almost all τ" we obtain a definition of a weakly admissible curve.

If a function $F(t)$ is defined and summable over a set $E - N$, where E is measurable and N has measure 0, we shall define

$$\int_E F(t)\,dt = \int_{E-N} F(t)\,dt.$$

This allows us to write, for example,

$$\int_a^b x'(t)\,dt = x(b) - x(a)$$

if $x(t)$ is absolutely continuous on $[a, b]$.

2. We now can state:

THEOREM I. *Under the hypotheses of* § 1, *if there exists an admissible curve, then there is an admissible curve* C *traversable in a time interval* $[a, a + T]$ *for which the time of traversal* T *is the least possible.*

Let T_0 be the greatest lower bound of the times of traversal of all admissible curves and let W be the lower bound of the times of traversal of all weakly admissible curves. Since the latter class contains the former, $W \leqq T_0$.

We now choose a sequence of weakly admissible curves C_n: $x = x_n(\tau)$, $0 \leqq \tau \leqq 1$, traversable in the respective intervals $[a_n, a_n + T_n]$, for which $T_n \to W$. From these we can select a subsequence such that a_n tends to a definite limit a_0; we suppose that $\{C_n\}$ is already such a sequence. Since Δ is closed, $a_0 \varepsilon \Delta$. All the time intervals $[a_n, a_n + T_n]$ lie in a bounded closed time interval, and all x lie in the bounded closed set A, and $V(x, t)$ is continuous; so for all such (x, t) the body $V(x, t)$ lies in a sphere about the origin of finite radius M. Then $|x'_n(\tau)|/T_n \leqq M$, so $|x'_n(\tau)|$ is uniformly bounded. Hence by Hilbert's theorem we can select a subsequence of the $x_n(\tau)$ which converges uniformly to a Lipschitzian limit function $x_0(\tau)$. We wish to prove that $x_0(\tau)$ is the curve sought. The proof will be given in a lemma.

LEMMA 2.1. *If the curves* C_n: $x = x_n(\tau)$, $0 \leqq \tau \leqq 1$ *are weakly admissible curves traversable in the intervals* $[a_n, a_n + T_n]$, *and*

(2.3) $\qquad a_n \to a_0, \quad T_n \to U, \quad x_n(\tau) \to x_0(\tau)$ *uniformly in* τ,

then C_0: $x = x_0(\tau)$ *is an admissible curve traversable in the time interval* $(a, a + U)$.

Condition (1.2a) clearly holds, for $x_0(t)$ has bounded derivatives as we saw just above. (1.2b) holds by the closure of A, and (1.2c) by the closure of Δ. (1.2d) is evident, for $x_0 = x_n(0) \to x_0(0)$ and $x_1 = x_n(1) \to x_0(1)$. We now turn to the proof of (1.2e).

Suppose that τ_0 is any number in $[0, 1]$ such that $x'_0(\tau)$ exists. If ϵ is a positive number, we define V_ϵ to be the set of all points whose distance from

$V(x_0(\tau_0),\ a_0 + U\tau_0)$ is $\leqq \epsilon$. By (2.3) and the continuity of $x_0(\tau)$ and $V(x,t)$, we find that

(2.4) *there exists a $\delta > 0$ and an integer n_0 such that if $|\tau - \tau_0| < \delta$ and $n > n_0$, then $V(x_n(\tau), a_n + T_n\tau) < V_\epsilon$.* Therefore, for all $n > n_0$ and almost all τ such that $|\tau - \tau_0| < \delta$,

$$x'_n(\tau) \in V_\epsilon. \tag{2.5}$$

Now suppose $0 < |h| < \delta$. Since V_ϵ is closed and convex, by Jensen's inequality [4]

$$(x_n(\tau_0 + h) - x_n(\tau_0))/h = (1/h) \int_{\tau_0}^{\tau_0+h} x'_n(\tau)\,d\tau \in V_\epsilon.$$

Let h be fixed and let $n \to \infty$. By the closure of V_ϵ,

$$[x_0(\tau_0 + h) - x_0(\tau_0)]/h = \lim_{n\to\infty} [x_n(\tau_0 + h) - x_n(\tau_0)]/h \in V_\epsilon.$$

Now let $h \to 0$. Again by the closure of V_ϵ, $x'_0(\tau_0) \in V_\epsilon$. But here the point $x'_0(\tau_0)$ does not depend on ϵ, and it can only belong to V_ϵ for every $\epsilon > 0$ if it belongs to $V_0 \equiv V(x_0(\tau_0), a_0 + U\tau_0)$. Hence (1.2e) is satisfied and Lemma 2.1 is established.[5]

Returning to the proof of Theorem I, by Lemma 2.1, the curve C_0 is an admissible curve traversable in time $W \leqq T_0$. But the time of traversal of any admissible curve is $\geqq T_0$. Hence the time of traversal of $C_0 = W = T_0$. This completes the proof of Theorem I.

Remark. We could alter the problem by assuming that $u(x,t)$ and $K(x,t)$ are defined only for t in a closed interval $t_0 \leqq t \leqq t_1$; nothing in the preceding demonstration would be altered.

From Lemma 2.1 we draw another conclusion:

LEMMA 2.2. *Every weakly admissible curve is admissible.*

For let $C: x = x(\tau)$, $0 \leqq \tau \leqq 1$ be a weakly admissible curve traversable in the interval $(a, a + T)$. In Lemma 2.1 we take $x_n(\tau) = x_0(\tau) = x(\tau)$, $T_n = T = U$ for all n. Then the hypotheses of the lemma are satisfied, and the conclusion informs us that the limit curve (which is C itself) is an admissible curve traversable in the interval $[a, a + T]$.

---

[4] Cf., for example, E. J. McShane, "On Jensen's inequality," *Bulletin of the American Mathematical Society*, vol. 40 (1937).

[5] We have proved a little more than we stated. If τ_0 is arbitrary, the above proof shows that every vector which is the limit of a sequence $[x_0(\tau_0 + h_m) - x_0(\tau_0)]/h_m$ as $h_m \to 0$ must belong to $V(x(\tau_0), a + T\tau_0)$, even though $x'_0(\tau_0)$ may not exist.

3. For the next theorem we shall add the hypothesis that the ship's engines are powerful enough so that at any time and place it can proceed with speed $\geqq \delta$ (δ a positive number) in any desired direction. Analytically, this means that the set of velocities which the ship can attain, namely

$$V(x,t) = u(x,t) + K(x,t),$$

shall contain all velocities v for which $|v| \leqq \delta$; that is, the sphere $|v| \leqq \delta$ is contained in $V(x,t)$ for all x in A and all t. If this is the case, for each direction (unit vector) d and each (x,t) there is just one number $\rho = \rho(d,x,t)$ such that ρd is on the boundary of $V(x,t)$; and $\rho \geqq \delta$ for all d, all x in A and all t. It is quite easy to see that if $\rho(d,x,t)$ is continuous, then the body $V(x,t)$ is a continuous function of (x,t). It is somewhat less easy to see that the converse is true;[6] so in order to save some space we shall hereafter replace the assumption that $V(x,t)$ is continuous by the assumption (only apparently stronger) that $\rho(d,x,t)$ is continuous. Then

THEOREM II. *Let the boundary of the body $V(x,t)$ be given in polar coördinates by the equation $\rho = \rho(d,x,t)$, where $\rho(d,x,t) \geqq \delta > 0$ and ρ is a continuous function of all its arguments. Then, if there exists an admissible curve, there is an admissible curve $C: x = X(t)$, $a \leqq t \leqq a+T$ such that the time T of traversal is the least possible, and such moreover that*

$$0 < |X'(t)| = \rho(X'(t)/|X'(t)|, x(t), t)$$

for almost all t.

By Theorem I, there is a curve $C: x = x(t)$, $a \leqq t \leqq a+T$ which can be traversed in time T which is the least possible time of traversal of any admissible curve. Let $x = \xi(s)$, $0 \leqq s \leqq L$ be the representation of C with arc length as parameter. To each t in $(a, a+T)$ there corresponds a value $s_0(t)$ of the parameter s, and $s'_0(t) = |x'(t)|$ for almost all t. This function is monotonic increasing and absolutely continuous. It may not have a single valued inverse, but we define $t_0(s)$ to be the least t such that $s_0(t) = s$. Then $t_0(s)$ is defined and single valued (possibly discontinuous) for $0 \leqq s \leqq L$, and $s_0(t_0(s)) \equiv s$. Also

$$t_0(0) = a, \qquad t_0(L) = a + T. \tag{3.1}$$

Since $x'(t) \in V(x(t),t)$, we see that $|x'(t)|$ is bounded, say $\leqq M$. Then for any times t_1, t_2 we find

$$|s_0(t_1) - s_0(t_2)| = \left| \int_{t_0}^{t_1} s'_0(t)\,dt \right| \leqq M \,|t_0 - t_1|,$$

[6] This follows, for example, from pages 14 and 37 of Bonnesen and Fenchel, *Theorie der konvexen Körper*.

whence for all values $s_1, s_2 > s_1$ of s

$$(3.2) \qquad t_0(s_2) - t_0(s_1) \geqq (s_2 - s_1)/M.$$

The derivative $t'_0(s)$ exists and is finite for almost all s. Inequality (3.2) shows $t'_0 \geqq M^{-1}$, so

(3.3) *for almost all s the derivatives $\xi'(s)$, $t'_0(s)$, and $s'_0(t_0(s))$ exist and are finite, and $|\xi'(s)| = 1$, and $s'_0(t)t'_0(s) = 1$.*

From this and the identity $x(t) = \xi(s_0(t))$ it follows that for almost all s the derivative $x'(t)$ exists and

$$(3.4) \qquad x'(t) = \xi'(s_0)s'_0(t).$$

Let S be the set of all s (of measure L) on which (3.3) and (3.4) hold. We recall that for almost all t (and, a fortiori, almost all s) $x'(t) \varepsilon V(x(t), t)$, so almost everywhere in S

$$(3.5) \quad 0 < s'_0(t_0(s)) = |x'(t_0(s))| \leqq \rho(x'(t_0(s))/|x'(t_0(s))|, x(t_0(s)), t_0(s)) \\ = \rho(\xi'(s), \xi(s), t_0(s)).$$

This implies that almost everywhere in S

$$(3.6) \qquad t'_0(s) = [s'_0(t_0(s))]^{-1} \geqq [\rho(\xi'(s), \xi(s), t_0(s))]^{-1}.$$

The function $\rho(\xi'(s), \xi(s), t)$ is defined for almost all s, is measurable in s for fixed t (being a continuous function of the measurable functions $\xi'(s)$, $\xi(s)$) and is continuous in t for fixed s. Moreover $\rho \geqq \delta$. Hence for every $\epsilon > 0$ the function $[\rho(\xi'(s), \xi(s), t) + \epsilon]^{-1}$ is measurable in s for fixed t, continuous in t for fixed s, and bounded. Therefore the equation

$$(3.7) \qquad t_\epsilon(s) = a - \epsilon + \int_0^s [\rho(\xi'(s), \xi(s), t_\epsilon(s)) + \epsilon]^{-1} ds$$

has an absolutely continuous solution $t_\epsilon(s)$ on $(0, L)$. We now prove

(3.8) *If $0 \leqq \beta < \epsilon$, then $t_\beta(s) > t_\epsilon(s)$ $(0 \leqq s \leqq L)$.*

The graphs of $t = t_\beta(s)$ and $t = t_\epsilon(s)$ are continuous curves. The latter is obvious, for $t_\epsilon(s)$ is continuous. So is the former if $\beta > 0$. If $\beta = 0$, the graph can still be considered as the continuous curve $s = s_0(t)$, $a \leqq t \leqq a + T$. It is clear that (3.8) holds at $s = 0$, for

$$t_\beta(0) - t_\epsilon(0) = (a - \beta) - (a - \epsilon) > 0.$$

If (3.8) is not always true, the graph of $t = t_\epsilon(s)$ lies somewhere above the graph of $t = t_\beta(s)$, and so they must have a first intersection point (σ, τ), $s_\epsilon(\tau) = s_\beta(\tau) = \sigma$. (Here $s_\epsilon(\tau)$ is the inverse of $t_\epsilon(s)$.) Thus for $s < \sigma$

$$t_\beta(s) > t_\epsilon(s). \tag{3.9}$$

By the uniform continuity of ρ, there is an $h > 0$ such that

$$\rho(\xi'(s), \xi(s), t_1) < \rho(\xi'(s), \xi(s), t_2) + \epsilon - \beta \text{ if } | t_1 - t_2 | < h. \tag{3.10}$$

By the continuity of $t_\epsilon(s)$, there is a $k > 0$ such that

$$0 \leqq t_\epsilon(\sigma) - t_\epsilon(s) < h \text{ if } \sigma - k \leqq s \leqq \sigma. \tag{3.11}$$

So for $\sigma - k \leqq s < \sigma$ we have, by (3.9) and (3.11),

$$\tau - h = t_\epsilon(\sigma) - h \leqq t_\epsilon(s) < t_\beta(s) \leqq \tau,$$

and therefore, by (3.10), for almost all s in $(\sigma - k, \sigma)$

$$\rho(\xi'(s), \xi(s), t_\beta(s)) + \beta < \rho(\xi'(s), \xi(s), t_\epsilon(s)) + \epsilon. \tag{3.12}$$

By (3.12) and (3.7) (and (3.6) if $\beta = 0$)

$$\tau'_\beta(s) > \tau'_\epsilon(s) \tag{3.13}$$

for almost all s in $(\sigma - k, \sigma)$. Now τ_β is monotonic increasing and τ_ϵ is absolutely continuous, so,[7] by (3.13)

$$\tau_\beta(\sigma) - \tau_\epsilon(\sigma) \geqq \tau_\beta(\sigma - k) + \int_{\sigma-k}^{\sigma} \tau'_\beta(s)\,ds - \tau_\epsilon(\sigma - k) - \int_{\sigma-k}^{\sigma} \tau'_\epsilon(s)\,ds$$
$$> \tau_\beta(\sigma - k) - \tau_\epsilon(\sigma - k) > 0.$$

Since τ_ϵ is continuous and τ_β is monotonic increasing, we can find an interval $\sigma \leqq s \leqq \sigma + l$ on which

$$\tau_\beta(s) - \tau_\epsilon(s) \geqq \tau_\beta(\sigma) - \tau_\epsilon(s) > 0.$$

So the graph of τ_β lies above the graph of τ_ϵ for $0 \leqq s \leqq \sigma + l$, and (σ, τ) cannot be an intersection point. This contradiction establishes statement (3.8).

Now let ϵ tend to 0 through a monotonic decreasing sequence of values. By (3.8) the successive functional values $t_\epsilon(s)$ increase for each s, so there exists a limit function:

$$\lim_{\epsilon \to 0} t_\epsilon(s) = \tau(s), \qquad 0 \leqq s \leqq L. \tag{3.14}$$

[7] Hobson, *Theory of Functions of a Real Variable*, vol. I, p. 590.

Also by (3. 8), $t_\epsilon(s) \leqq t_0(s)$ for each $\epsilon > 0$, so in the limit

$$(3.15) \qquad \tau(s) \leqq t_0(s), \qquad 0 \leqq s \leqq L.$$

In equation (3. 7), the integrand on the right is uniformly bounded and tends almost everywhere to $[\rho(\xi'(s), \xi(s), \tau(s))]^{-1}$, while on the left $t_\epsilon \to \tau$; so

$$(3.16) \qquad \tau(s) = a + \int_0^s [\rho(\xi'(s), \xi(s), \tau(s))]^{-1} ds.$$

The function $\tau(s)$ has almost everywhere a positive derivative, by (3. 16), so it has an absolutely continuous inverse $s = s(\tau)$, $a \leqq \tau \leqq \tau(L)$. Defining $X(\tau) = \xi(s(\tau))$, we find that for almost all τ

$$X'(\tau) = \xi'(s)s'(\tau), \qquad s'(\tau) > 0,$$

so $X'(\tau)/|X'(\tau)| = \xi'(s)$ for almost all τ. Moreover, for almost all τ

$$\begin{aligned} X'(\tau) &= \xi'(s)[\tau'(s)]^{-1} = \xi'(s)\rho(\xi'(s), \xi(s), \tau(s)) \\ &= \xi'(s(\tau))\rho(X'(\tau)/|X'(\tau)|, X(\tau), \tau), \end{aligned}$$

whence

$$|X'(\tau)| = \rho(X'(\tau)/|X'(\tau)|, X(\tau), \tau)$$

for almost all τ. That is, condition (2. 1e) holds with the words "wherever $X'(\tau)$ is defined" replaced by "for almost all τ." The others of conditions (2. 1) are trivially easily verified, so $x = X(\tau)$ is weakly admissible, and by Lemma 2. 2 it is admissible. All that remains to prove is that the time of traversal $\tau(L) - a$ has the least possible value T. Clearly $\tau(L) - a \geqq T$. By (3. 15) and (3. 1), $\tau(L) \leqq t_0(L) = a + T$, completing the proof.

Remark. If we change the problem by assuming that $u(x, t)$ and $K(x, t)$ are defined only on a time interval $t_0 \leqq t \leqq t_1$, one point of the preceding demonstration needs change. In defining t_ϵ, times $t < a$ entered. If however we define $V(x, t) = V(x, t_0)$ for $t < t_0$, the proof can be carried out as above, and the final results will involve only times t in $[t_0, t_1]$.

UNIVERSITY OF VIRGINIA.

THE TOPOLOGICAL DISCRIMINANT GROUP OF A RIEMANN SURFACE OF GENUS p.

By Oscar Zariski.

1. Introduction. The symmetric n-th product K^n of a complex K carries a subset D whose points represent n-tuples on K with two or more coincident points. We call D the *discriminant variety* of K^n and we refer to the fundamental group of the residual space $K^n - D$ as the *topological discriminant group* (*of degree* n) of the given complex K. In part I we determine this group, $G_{n,p}$, when K is a Riemann surface R of genus p. We were led to examine this group by the following considerations. The variety R^n is the space of all n-tuples of points of an algebraic curve f of genus p. As such, R^n carries—for a sufficiently high value of n ($n \geqq p+2$)—a system, ∞^p, of linear $(n-p)$-spaces S_{n-p}, images of complete linear series $g_n{}^{n-p}$ existent on f. If D_1 denotes the intersection of a general S_{n-p} of the system with the discriminant variety D of R^n, the fundamental group of the residual space $S_{n-p} - D_1$ can be shown to be an invariant subgroup $H_{n,p}$ of $G_{n,p}$, and the quotient group $G_{n,p}/H_{n,p}$ is simply isomorphic to the homology group of R. If $G_{n,p}$ is known, $H_{n,p}$ can be determined on the basis of well-known principles laid down by Reidemeister.[1] By a theorem which we have proved elsewhere (Zariski [5]) the fundamental group of $S_{n-p} - D_1$ coincides with the fundamental group of the residual space of a general plane section C of D_1. It is not difficult to see that C is the plane dual of a general plane curve of order n and genus p, so that C is of order $2n+2p-2$ with $3(n+2p-2)$ cusps and $2(n-2)(n-3)+2p(2n+p-7)$ nodes. The knowledge of the fundamental group of C, of interest in itself, makes it also possible to determine the fundamental group of any plane curve admitting C as a limiting case. In this connection we may point out that the class of curves thus obtained is not negligible, since, at present, duality constructions and limiting processes are, with a few exceptions, the only means of arriving at effectively existent curves with nodes and cusps.

In Parts II and III we carry out in detail the above outlined considerations in the case $p = 1$. The somewhat elaborate group-theoretic apparatus of Part II is inherent to the reduction of the infinite set of generators and generating relations of $H_{n,1}$ (an invariant subgroup of $G_{n,1}$ of infinite index) to a finite set of generators and generating relations. The existence of such a

finite set is, *a priori*, implied by the algebro-geometric interpretation of $H_{n,1}$ given in Part III.

An interesting special case, examined in Part III, is given by the dual of a plane cubic—a sextic with 9 cusps. It is then found that the 9 generating relations at the cusps enjoy properties which are in striking analogy with the well-known alignment properties of the configuration of the 9 flexes of a plane cubic.

2. Let R be a Riemann surface of genus p and let R^n be the symmetric n-th product of R, i. e. the space (of n complex dimensions) of all unordered n-tuples of points of R, topologized in an obvious manner. We have shown elsewhere (Zariski,[3] p. 1), that R^n is a manifold. We denote by D the subvariety of R^n whose points correspond to n-tuples of points of R in which two or more points coincide. This variety D—which can legitimately be designated as the *discriminant variety of* R^n—is of $n-1$ complex dimensions. The purpose of this and of the next two sections is the determination of the fundamental group of the residual space $R^n - D$. We denote this group by $G_{n,p}$.

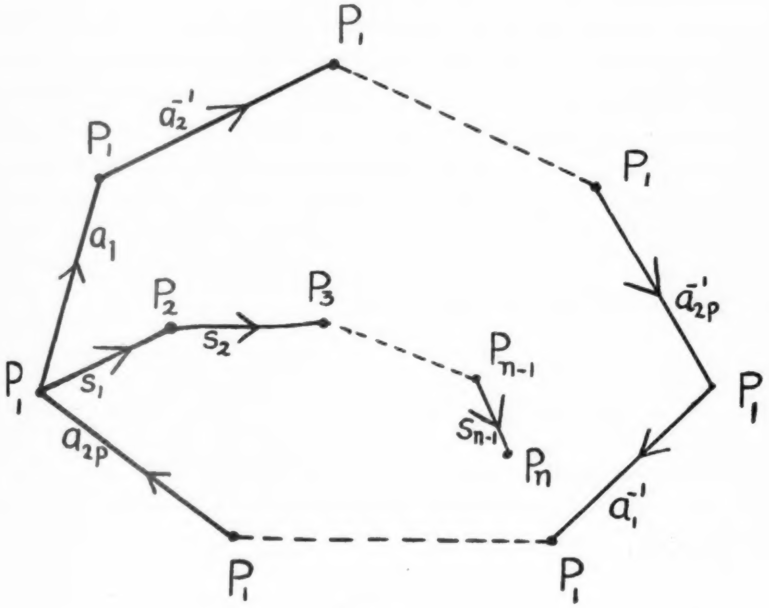
The group $G_{n,0}$ is known (see Zariski,[4] p. 612). As in the just quoted paper, we interpret also here the group $G_{n,p}$ as the *group of motion classes of* n *points of* R. The motions considered are those which carry a fixed initial set of n distinct points $P_1, P_2, \cdots, P_n$ of R into its initial position (allowing for a permutation of the points P_i) and in the course of which the variable set consists always of distinct points. Two motions belong to the same class if they can be deformed into each other through a continuous chain of motions of the same nature.

We fix on R a set of retrosections $a_1, a_2, \cdots, a_{2p}$ on a common point P_1, belonging to our initial n-tuple of points. We choose our retrosections in such a manner that when R is cut open along them, the resulting 2-cell E_2 is bounded by the closed polygon

$$a_1 a_2^{-1} a_3 \cdots a_{2p}^{-1} a_1^{-1} a_2 a_3^{-1} \cdots a_{2p}.$$

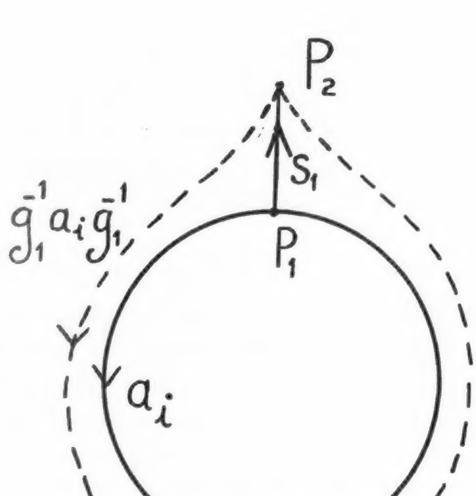
We assume the points $P_2, \cdots, P_n$ in the interior of E_2 and we join the points $P_1, P_2, \cdots, P_n$ by a set of simple oriented arcs $s_1, s_2, \cdots, s_{n-1}$ (see figure 1). The indicated orientations of the retrosections a_i and of the arcs s_j are such that at P_1 the positive sense on each retrosection a_i points from the right-hand edge of the arc s_1 toward its left-hand edge (see figure 2).

We denote by g_i the motion in which P_i is carried into P_{i+1} along the left-hand edge of s_i and P_{i+1} is carried into P_i along the right-hand edge of s_i, while the remaining points P_k are fixed in their initial positions. In the case



P_1
a_2^{-1}
P_1
P_1
a_1
P_2
P_3
a_{2p}^{-1}
s_2
P_{n-1}
P_1
s_1
P_1
s_{n-1}
a_{2p}
P_n
a_1^{-1}
P_1
P_1

FIG. 1.



P_2
s_1
$g_1^{-1} a_i g_1^{-1}$
P_1
a_i

FIG. 2.

$p=0$ the elements $g_1, g_2, \cdots, g_{n-1}$ are generators of $G_{n,0}$ (see Zariski,[4] p. 610). It follows that any motion in the course of which the points of the variable set do not cross the boundary of E_2 can be expressed as a product of the g_i's. Let us consider the motion in which P_1 describes the oriented retrosection a_i, while the remaining points P_k are fixed ($k>1$). We shall denote this motion by the same letter a_i. It is obvious that any crossing of the boundary of E_2 introduces factors $a_i^{\pm 1}$, hence the *elements* $g_1, \cdots, g_{n-1}, a_1, \cdots, a_{2p}$ *are the generators of* $G_{n,p}$.

3. The generating relations of $G_{n,p}$. The relations

$$(\alpha) \qquad g_i g_j = g_j g_i, \qquad |i-j| \neq 1;$$

$$(\beta) \qquad g_i g_{i+1} g_i = g_{i+1} g_i g_{i+1}, \qquad (i=1,2,\cdots,n-2);$$

established in our quoted paper (Zariski,[4] p. 612) remain valid also in the present case. The relation (6) of the quoted paper now has to be replaced by the following:

$$(\gamma) \qquad g_1 \cdots g_{n-2} g^2_{n-1} g_{n-2} \cdots g_1 = a_1 a_2^{-1} a_3 \cdots a_{2p}^{-1} a_1^{-1} a_2 \cdots a_{2p},$$

since the left-hand member represents a motion which, as is easily seen, can be deformed into one in which the points $P_2, \cdots, P_n$ are fixed, while P_1 describes a closed path surrounding the set $\{P_2, \cdots, P_n\}$. This closed curve can be deformed into the boundary of the cell E_2. Other generating relations, involving the g's and the elements a_i, are obtained as follows:

In the first place it is clear that each a_i is permutable with each of the elements $g_2, \cdots, g_{n-1}$, since the corresponding paths do not intersect. Hence

$$(\delta) \qquad g_k a_i = a_i g_k, \qquad (i=1,2,\cdots,2p;\ k=2,\cdots,n-1).$$

The motion $g_1^{-1} a_i g_1^{-1}$ can be deformed into a motion in which the points $P_1, P_3, \cdots, P_n$ are fixed, while P_2 describes a retrosection homologous to a_i and not meeting a_i (see figure 2, the path of P_2 is indicated by the punctuated curve). Hence a_i and $g_1^{-1} a_i g_1^{-1}$ are permutable, whence the relation

$$(\epsilon) \qquad (g_1^{-1} a_i)^2 = (a_i g_1^{-1})^2.$$

We now introduce the following elements:

$$(\mu) \qquad a'_i = g_1^{-1} a_i g_1.$$

It is clear that the motion a'_i is equivalent to a motion in which $P_1, P_3, \cdots, P_n$ are fixed, while P_2 describes a retrosection a'_i homologous to a_i, and that a'_i

and a_j intersect in one point only, provided $j > i$ (see figure 3, illustrating the behavior of a'_2). If we then consider the direct product of a'_i and a_j, regarded as 1-spheres, we have a torus T, on which the motions a_j and a'_i, regarded as motions of the point pair P_1P_2 (the remaining points $P_3, \cdots, P_n$ being fixed), are represented by two retrosections α and α' respectively. The common point $\bar{P}$ of α' and α corresponds to the initial point-pair (P_1, P_2). Let Q be the point at which a'_i and a_j intersect, and let $\bar{Q}$ be the point of the torus which corresponds to the point pair (Q, Q). A closed path on $T - \bar{Q}$ starting from and returning to $\bar{P}$ represents a motion of a variable pair of

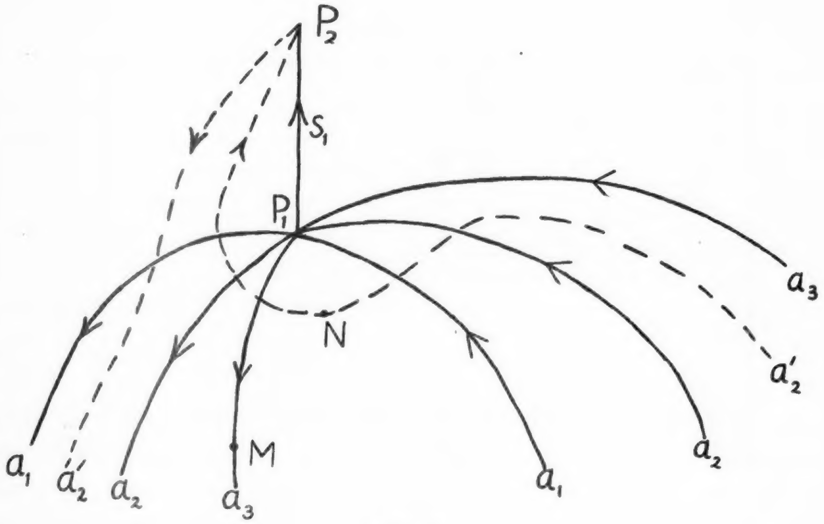


FIG. 3.

distinct points starting from and returning to the initial point pair P_1, P_2. A deformation of this path on $T - \bar{Q}$ corresponds to an allowable deformation of the corresponding motion on R. Since on $T - \bar{Q}$ we have $\alpha\alpha'^{-1}\alpha^{-1}\alpha' = \bar{\gamma}$, where $\bar{\gamma}$ is a properly oriented loop issued from $\bar{P}$ and surrounding the point $\bar{Q}$, we have a corresponding relation $a_j a'^{-1}_i a_j^{-1} a'_i = \gamma$, where γ is the motion of the variable point pair on R which corresponds to the loop $\bar{\gamma}$ on T. To determine γ, we take as $\bar{\gamma}$ a quadrangle two of whose sides are on the retrosections α and α' and the other two are parallel to these retrosections (see Fig. 4). The corresponding motion γ has now the following description: (a) first the point P_1 describes the arc P_1M on a_j, P_2 is fixed; (b) then P_2 describes the arc P_2N on a'_i, while the second point is fixed at M; (c) a reversal of motion (a); (d) the reversal of motion (b) (see Fig. 3, where $j = 3$, $i = 2$).

By letting M approach P_1 on a_j and by accompanying this by a deformation of the path (b), we see immediately that the combined motion γ can be deformed into a motion in which P_1 is fixed and in which P_2 turns around P_1 in what on Fig. 3 would be the clockwise sense. This motion is visibly equivalent to the motion $g_1{}^2$. We have therefore the following generating relation:

$$(\nu) \qquad a_j a'_i{}^{-1} a_j{}^{-1} a'_i = g_1{}^2, \qquad j > i.$$

We prove in the next section that *the relations* (α), (β), (γ), (δ), (ϵ) *and* (ν)

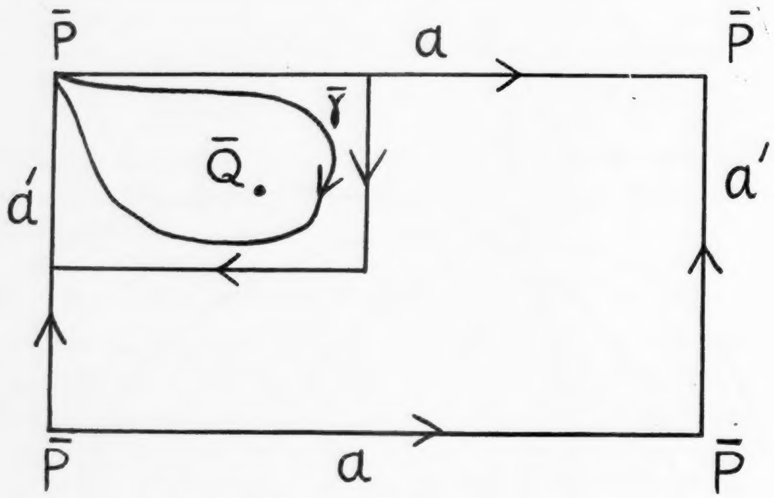


FIG. 4.

(where the elements a'_i are defined by (μ)) *constitute a complete set of generating relations of* $G_{n,p}$.

4. We denote the abstract group defined by the relations $(\alpha - \nu)$ of the preceding section by $\bar{G}_{n,p}$, and we use the notations $\equiv$ and $=$ to indicate equality of elements in $\bar{G}_{n,p}$ and $G_{n,p}$ respectively. We wish to prove that $\bar{G}_{n,p}$ coincides with $G_{n,p}$, or what is the same that $\alpha = \beta$ implies $\alpha \equiv \beta$, where α and and β are products of the generators a_i, g_k. The proof will be made in several steps.

(a) *If W is any product of the generators a_i, g_k, then $W \equiv g_h g_{h-1} \cdots g_1 W_1$, where $0 \leq h \leq n-1$ and where W_1 involves only the elements $g_2, \cdots, g_n, a_i, a'_i$ and the elements*

$$s_j = (g_{j-2}g_{j-3} \cdots g_1)^{-1} g^2_{j-1} (g_{j-2}g_{j-3} \cdots g_1), \qquad (j = 2, \cdots, n).$$

The proof is the same as the one given in Zariski,[4] pp. 612-613, except that it is also necessary to make use of the relations

$$a_i^{\pm 1} g_h \cdots g_1 = g_h \cdots g_2 a_i^{\pm 1} g_1 = g_h \cdots g_2 g_1 a'^{\pm 1}_i.$$

The elements s_j, denoted in Zariski,[4] p. 612, by a_j, represent motions in which the points $P_2, \cdots, P_n$ are at rest while P_1 describes a loop around the point P_j.

(b) *If W represents a motion in which P_1 returns to its initial position, then* $W \equiv W_1$.

This is a consequence of (a) since in $g_h \cdots g_1 W_1$ the point P_{h+1} is carried into P_1.

(c) *The subgroup Γ of $\bar{G}_{n,p}$ generated by the elements $s_2, \cdots, s_n, a_1, \cdots, a_{2p}$ is an invariant subgroup of the group generated in $\bar{G}_{n,p}$ by the above elements and by the elements $g_2, \cdots, g_{n-1}, a'_i$.*

That $g_k^{\pm 1} s_j g_k^{\mp 1} \subset \Gamma$ for $k = 2, \cdots, n-1$ has already been proved in Zariski,[4] p. 613. In view of (δ) it remains to prove that $a'^{\pm 1}_i \Gamma a'^{\mp 1}_i \subset \Gamma$.

The relation (ϵ) shows that a'_i is commutative with $s_2^{-1} a_i$ $(= g_1^{-2} a_i)$. Hence either one of the two relations

$$(1) \qquad a'^{\epsilon}_i s_2 a'^{-\epsilon}_i \subset \Gamma, \qquad a'^{\epsilon}_i a_i a'^{-\epsilon}_i \subset \Gamma, \qquad \epsilon = \pm 1$$

implies the other. Now (ϵ) also implies the relation $(a'_i)^{-1} a_i a'_i = s_2^{-1} a_i s_2$. Consequently both relations (1) hold true for $\epsilon = -1$. Transforming the relation $(a'_i)^{-1} a_i a'_i = s_2^{-1} a_i s_2$ by $(a'_i)^{-1}$ and taking into account the commutativity of the elements a'_i, $s_2^{-1} a_i$, we find that $a'_i s_2 a'^{-1}_i$ belongs to Γ. Hence the relations (1) hold true for $\epsilon = \pm 1$.

Since $g_1 s_j g_1^{-1}$, $j > 2$, involves only the elements $g_2, \cdots, g_{n-1}$, it follows by (δ) that $g_1 s_j g_1^{-1}$ is commutative with each a_i, and hence s_j is commutative with a'_i $(= g_1^{-1} a_i g_1)$, for $j > 2$.

It remains to prove that all the transforms $(a'_i)^{\epsilon} a_j (a'_i)^{-\epsilon}$, $\epsilon = \pm 1$, $i \neq j$, belong to Γ. For $\epsilon = -1$ and $i < j$ this follows directly from the relation (ν), and for $\epsilon = +1$ and $i < j$ this is proved by transforming the relation (ν) by $(a'_i)^{-1}$, since we have already proved that $a'_i g_1^2 a'^{-1}_i$ is in Γ. We now transform (ν) by g_1 and we obtain the relation $a'_j s_2^{-1} a_i s_2 a'^{-1}_j = s_2 a_i^{-1}$. Since $a'^{\epsilon}_j s_2 a'^{-\epsilon}_j \subset \Gamma$, for $\epsilon = \pm 1$, it follows immediately that $a'^{\epsilon}_j a_i a'^{-\epsilon}_j \subset \Gamma$, and this completes the proof of the invariance of the subgroup Γ.

(d) As a consequence of (b) and (c) we may now assert that if W represents a motion in which the point P_1 returns to its initial position, then we have already in $\bar{G}_{n,p}$ a relation of the form:

$$(2) \quad W \equiv W_1(g_2, \cdots, g_{n-1};\ a'_1, \cdots, a'_{2p}) \cdot W_2(s_2, \cdots, s_{n-1};\ a_1, \cdots, a_{2p}),$$

where W_1 and W_2 are products of the elements indicated in the parentheses.

(e) To prove that the groups $G_{n,p}$ and $\bar{G}_{n,p}$ are identical, we use an induction with respect to n, since for $n = 1$ the group $G_{n,p}$ is merely the fundamental group of the Riemann surface R, and in this case the relation (ν), where now the left-hand member reduces to 1, is the only generating relation for $G_{n,p}$. Hence $G_{1,p}$ coincides with $\bar{G}_{1,p}$. We shall then assume that $G_{n-1,p}$ and $\bar{G}_{n-1,p}$ are identical groups. For $G_{n-1,p}$ we take as initial sets of $n-1$ points the points $P_2, \cdots, P_n$, and as generators the elements $g_2, \cdots, g_{n-1}$, $a'_1, \cdots, a'_{2p}$. As elements analogous to $g_1^{-1}a_ig_1$ $(= a'_i)$ we take the elements $a''_i = g_2^{-1}a'_ig_2$.

Let W be an element of $G_{n,p}$, expressed as a product of the generators g_k, a_i, and let $W = 1$ be a true relation in $G_{n,p}$. Since in the motion W every point P_i returns to its initial position, the representation (2) of W holds true. Since in the motion W_1 the point P_1 is fixed, while in the motion W_2 the points $P_2, \cdots, P_n$ are fixed, it is clear that $W_1 = 1$ must be a true relation in $G_{n-1,p}$. By our induction, this relation must be a consequence of the relations $(\alpha - \nu)$ for the case $n - 1$. To rewrite these relations for the group $G_{n-1,p}$ we must replace $g_1, \cdots, g_n$ by $g_2, \cdots, g_n$ and the elements a_i by the elements a'_i. Let us letter these new relations by $\alpha', \beta', \cdots, \nu'$.

We assert that the relations (α'), (β'), (δ'), (ϵ'), (ν') *are group-theoretic consequences of the relations* (α), (β), (δ), (ϵ), (ν). The relations (α'), (β') are among the relations (α), (β). As for the relations (δ'), we observe that by (α), g_1 is commutative with g_k, $k \geq 3$, and hence the relations (δ') are obtained by transforming by g_1 those relations (δ) in which $k \geq 3$. Finally, the relations (ϵ'), (ν') are the transforms of (ϵ) and (ν) by g_2g_1. In fact, since g_2 and a_i are commutative, we have

$$(g_2g_1)^{-1}a_i(g_2g_1) = g_1^{-1}a_ig_1 = a'_i.$$

Taking into account the relation $g_1g_2g_1 = g_2g_1g_2$ we find

$$(g_2g_1)^{-1}a'_i(g_2g_1) = (g_1g_2)^{-1}a_i(g_1g_2) = g_2^{-1}a'_ig_2 = a''_i.$$

Moreover, we have $(g_2g_1)^{-1}g_1(g_2g_1) = g_2$, as a consequence of the relation $g_1g_2g_1 = g_2g_1g_2$.

We have left out the relation (γ'), i. e. the following:

$$H = a'_1a'^{-1}_2 \cdots a'^{-1}_{2p}a'^{-1}_1a'_2 \cdots a'_{2p}(g_2 \cdots g^2_{n-1} \cdots g_2)^{-1} = 1.$$

This is not a true relation in $\bar{G}_{n,p}$. In fact, if we transform (γ) by g_1, we find the following relation:

$$H \equiv g_1{}^2; \quad \text{i.e.,} \quad H \equiv s_2.$$

Having thus proved that all the generating relations of $G_{n-1,p}$, except the relation $H = 1$, are also true relations in $\bar{G}_{n,p}$, and recalling that $W_1 = 1$ holds in $G_{n-1,p}$ we deduce that W_1, as an element of the group $\bar{G}_{n,p}$, can be expressed as a product of transforms of H, i. e. of s_2, the transforming elements involving only the generators $g_2, \cdots, g_n, a'_i$ of $G_{n-1,p}$. Hence, by (c), we can write $W_1 \equiv W'_1$, where $W'_1 \subset \Gamma$. In view of (2), we conclude that *if a relation* $W = 1$ *holds true in* $G_{n,p}$, *then we have in* $\bar{G}_{n,p}$ $W \equiv F(a_1, \cdots, a_{2p}; s_2, \cdots, s_n)$, *where* F *is a product involving only the element* a_i, s_j.

The elements $a_1, \cdots, a_{2p}, s_2, \cdots, s_n$ are generators of the fundamental group G^* of the Riemann surface with $n - 1$ holes at $P_2, \cdots, P_n$, i. e. of $R - P_2 - \cdots - P_n$. It can be proved, as in Zariski,[4] p. 614, that the relation $F = 1$, true in $G_{n,p}$, implies that F, considered as an element of G^*, belongs to the center of G^*. Since G^* is a free group, for $n \geqq 2$, it follows that *F is the identity in G^**. The only generating relation of G^* is the following:

$$a_1 a_2{}^{-1} \cdots {}^{-1}_{2p} a_1{}^{-1} a_2 \cdots a_{2p} = s_2 s_3 \cdots s_n.$$

If the s_j's are replaced by their expressions in terms of the g_k's, this relation coincides with the relation (γ). Hence $F = 1$ is also a true relation in $\bar{G}_{n,p}$, i. e. we have $F \equiv 1$. Consequently $W = 1$ implies $W \equiv 1$, q. e. d.

II. On an invariant subgroup of $G_{n,p}$ in the elliptic case.

5. In a motion which carries the initial n-tuple $P_1, \cdots, P_n$ back to its initial position, the paths described by the points $P_1, \cdots, P_n$ constitute together a closed curve, a singular 1-cycle σ_1. Those elements of $G_{n,p}$ for which this cycle σ_1 is ~ 0 on the Riemann surface R form an invariant subgroup $H_{n,p}$ of $G_{n,p}$, and the quotient group $G_{n,p}/H_{n,p}$ is the homology group of R. There is a general procedure, given by Reidemeister,[1] for determining the generators and the generating relations (finite or infinite in number) of an invariant subgroup of a discrete infinite group, whether the quotient group is finite or infinite. We shall now apply this general method to the invariant group $H_{n,p}$ of $G_{n,p}$ in the elliptic case ($p = 1$). It will be seen that $H_{n,1}$ admits a finite set of generators satisfying a finite set of relations, although the quotient group is in this case a free abelian group. This could also be foreseen from the geometric considerations of Part III of this paper, where it will be shown that the group $H_{n,1}$ is the fundamental group of the residual space of a certain elliptic plane curve.

The reduction of the set of generators and of generating relations of $H_{n,1}$, which we are about to undertake, can be extended to the group $H_{n,p}$, p arbitrary, in so far at least as the explicit determination of a finite set of generators is concerned. As for the generating relations, a similar reduction presents some difficulties.

From now on we shall denote the groups $G_{n,1}$ and $H_{n,1}$ by G_n and H_n respectively. G_n is generated by the elements $a_1, a_2, g_1, \cdots, g_{n-1}$. We rewrite the generating relations of G_n as follows:

$$(3)\qquad \begin{cases} T_{ij} = g_i g_j g_i^{-1} g_j^{-1} = 1, & |i-j| \neq 1; \\ T_{i,i+1} = g_i g_{i+1} g_i g_{i+1}^{-1} g_i^{-1} g_{i+1}^{-1} = 1; & \\ T = g_1 \cdots g_{n-2} g^2_{n-1} g_{n-2} \cdots g_1 a_2^{-1} a_1 a_2 a_1^{-1} = 1; & \\ P_{ik} = a_i g_k a_i^{-1} g_k^{-1} = 1, & (k = 2, \cdots, n-1;\ i = 1, 2), \\ S_i = (a_i g_1^{-1})^2 (g_1^{-1} a_i)^{-2} = 1, & (i = 1, 2); \\ S_{12} = a_2 a'^{-1}_1 a_2^{-1} a'_1 g_1^{-2} = 1, \quad a'_1 = g_1^{-1} a_1 g_1. & \end{cases}$$

Since a_1, a_2, considered as 1-cycles on R, are generators of the homology group, it follows that there is a $(1-1)$ correspondence between the elements $a_1^k a_2^l$ of G_n and the elements of the quotient group G_n/H_n. It is also clear that the elements of H_n are those and only those elements of G_n which become equal to 1 if the relations $g_i = 1$, $a_1^{-1}a_2a_1a_2^{-1} = 1$ are added to the generating relations of G_n, i. e. those power products of the generators a_1, a_2, g_k in which the sum of the exponents of each of the elements a_1, a_2 vanishes. It follows, by the quoted paper of Reidemeister, that the following elements are generators of H_n:

$$(4)\qquad \begin{cases} \alpha_{ij} = (a_1^i a_2^j) a_1 (a_1^{i+1} a_2^j)^{-1}, & \\ \alpha'_{ij} = (a_1^i a_2^j) a_2 (a_1^i a_2^{j+1})^{-1}, & (i, j = 0, \pm 1, \pm 2, \cdots). \\ g_{ij} = (a_1^i a_2^j) g_1 (a_1^i a_2^j)^{-1}, & \\ g_{ij,k} = (a_1^i a_2^j) g_k (a_1^i a_2^j)^{-1}, & (k = 2, \cdots, n-1). \end{cases}$$

The elements α'_{ij} are all identically equal to 1, and hence we are left with the generators $\alpha_{ij}, g_{ij}, g_{ij,2}, \cdots, g_{ij,n-1}$.

Generating relations of H_n are obtained as follows. In the first place a *definite construction* is given by means of which any power product π of the generators of G_n can be expressed in the form $\pi_1 a_1^i a_2^j$, where π_1 is a power product of the generators of H_n. Here the exponents i, j are uniquely determined by π, since π and $a_1^i a_2^j$ correspond to one and the same element of the quotient group G_n/H_n. This construction is as follows. Let $\pi = \pi' \lambda^{\pm 1}$, where λ is a generator of G_n and where π' contains less factors than π, and let us assume that we have already expressed π' in the form $\pi'_1 a_1^i a_2^j$. Among the generators (4) of H_n there is an element λ_{ij} of the form $(a_1^i a_2^j)\lambda(a_1^{i'} a_2^{j'})^{-1}$

and there is also an element $\bar{\lambda}_{ij}$ of the form $(a_1{}^{i''}a_2{}^{j''})\lambda(a_1{}^{i}a_2{}^{j})^{-1}$. Then, if $\pi = \pi'\lambda$, we write $\pi = \pi'_1\lambda_{ij}a_1{}^{i'}a_2{}^{j'}$ and if $\pi = \pi'\lambda^{-1}$ we write $\pi = \pi'_1\bar{\bar{\lambda}}_{ij}^{-1}a_1{}^{i''}a_2{}^{j''}$.

Using this construction, we obtain all the generating relations of H_n as follows: (a) we express the power products which occur in the relations (3) by means of the generators of H_n and we put equal to 1 the resulting expressions; (b) we apply the same procedure to the transforms of the relations (3) by any of the elements $a_1{}^{i}a_2{}^{j}$; (c) we finally apply our construction to the elements $a_1{}^{i}a_2{}^{j}$, getting $a_1{}^{i}a_2{}^{j} = \pi_{ij}a_1{}^{i}a_2{}^{j}$, where π_{ij} is a power product of the generators of H_n, and we put $\pi_{ij} = 1$.*

It is immediately seen that the relations $\pi_{ij} = 1$ give only the following trivial relations: †

$$\alpha'_{ij} = 1, \text{ for all } i \text{ and } j;$$
$$\alpha_{i0} = 1, \text{ for all } i.$$

The relations $(a_1{}^{i}a_2{}^{j})P_{ak}(a_1{}^{i}a_2{}^{j})^{-1}$ merely imply $g_{ij,k} = g_k$ and $\alpha_{ij}g_k = g_k\alpha_{ij}$ for all i, j and for $k \geq 2$.

Reassuming, we have at present the following generators and generating relations for the group H_n:

Generators of H_n:

$$\alpha_{ij},\, g_{ij} \quad (i, j = 0, \pm 1, \pm 2, \cdots); \qquad g_2, \cdots, g_{n-1}.$$

Generating relations of H_n:

---

* The relations $\pi_{ij} = 1$ replace in the present case the relations $T_mF_{kl}T_m{}^{-1}$ given by Reidemeister,[1] p. 13. We use the notations of Reidemeister and we prove that, quite generally, *the ng^2 relations $T_mF_{kl}T_m{}^{-1} = 1$ can be replaced by the g relations $\pi_m(S_{ik}) = 1$*, where $\pi_m(S_{ik})$ is the power product of the S_{ik}'s which we get if we express T_m in the form π_mT_m according to Reidemeister's construction. Let us first consider the case in which $m \neq g$, i. e. T_m is not the element 1. In this case $T_mF_{kl}T_m{}^{-1}$ contains S_{kl}^{-1} and hence (see Reidemeister,[1] p. 11) in the course of the construction S_{kl}^{-1} must be replaced by $T_{g_{ik}}S_k{}^{-1}T_i{}^{-1}$. But then we get an expression which can be changed into $T_mT_m{}^{-1}$ by using the trivial relations $S_iS_i{}^{-1} = 1$. Hence the relation $T_mF_{kl}T_m = 1$, expressed in terms of S_{ik}'s, can be changed into the relation $T_mT_m{}^{-1} = 1$ by using the trivial relations $S_{ik}S_{ik}^{-1} = 1$. Now it is immediately seen that this last relation coincides with the relation $\pi_m{}^{-1}(S_{ik}) = 1$. Let now $m = g$. In this case we have the generating relation $F_{kl} = 1$; but F_{kl} expressed in terms of the S_{ik}'s has the following form:

$$S_{kl}^{-1}\pi_k(S_{ij})S_{kl}\pi_{g_{kl}}^{-1}(S_{ij}),$$

and hence the relations $F_{kl} = 1$ are consequences of the relations $\pi_m = 1$.

† If, for instance, $i \geq 0$ and $j \geq 0$, then $\pi_{ij} = \alpha_{00}\alpha_{10} \cdots \alpha_{i-1,0}\alpha'_{i,0}\alpha'_{i1}, \alpha'_{i,j-1}$.

$$\left.\begin{array}{lll}
(5) & T^{(ij)}_{kl} = (a_1{}^i a_2{}^j) T_{kl} (a_1{}^i a_2{}^j)^{-1} = 1, & \\
 & \quad (k, l = 1, 2, \cdots, n-1, k \neq l) & \\
(5') & T^{(ij)} = (a_1{}^i a_2{}^j) T \;\; (a_1{}^i a_j{}^j)^{-1} = 1, & \\
(5'') & S_k{}^{(ij)} = (a_1{}^i a_2{}^j) S_k \;\; (a_1{}^i a_2{}^j)^{-1} = 1, & \\
 & \quad (k = 1, 2); & \\
(5''') & S^{(ij)}_{12} = (a_1{}^i a_2{}^j) S_{12} (a_1{}^i a_2{}^j)^{-1} = 1 &
\end{array}\right\} (i, j = 0, \pm 1, \pm 2, \cdots).$$

$$(5a) \qquad \alpha_{i0} = 1, \qquad (i = 0, \pm 1, \pm 2, \cdots),$$

$$(5b) \qquad \alpha_{ij} g_k = g_k \alpha_{ij}, \quad (k \geq 2).$$

6. The expression of the elements $T^{(ij)}_{kl}, T^{(ij)}, S_k{}^{(ij)}, S^{(ij)}_{12}$ in terms of the generators of H_n leads in a straightforward manner to the following generating relations (where $i, j = 0, \pm 1, \pm 2, \cdots$):

$$\begin{array}{lll}
(6) & T^{(ij)}_{1l}: & g_{ij} g_l = g_l g_{ij}, \qquad l > 2; \\
(6') & T^{(ij)}_{kl}: & g_k g_l = g_l g_k; \qquad | l - k | \neq 1; \; k, l > 1; \\
(6'') & T^{(ij)}_{12}: & g_{ij} g_2 g_{ij} = g_2 g_{ij} g_2; \\
(6''') & T^{(ij)}_{k,k+1}: & g_k g_{k+1} g_k = g_{k+1} g_k g_{k+1}, \qquad k > 1; \\
(7) & T^{(ij)}: & g_{ij} g_2 \cdots g_{n-2} g^2{}_{n-1} g_{n-2} \cdots g_2 g_{ij} = \alpha_{ij} \alpha^{-1}_{i,j-1}; \\
(8_1) & S_1{}^{(ij)}: & g_{i+2,j} = (\alpha_{ij} g^{-1}_{i+1,j} \alpha_{i+1,j})^{-1} g_{i,j} (\alpha_{ij} g^{-1}_{i+1,j} \alpha_{i+1,j}); \\
(8_2) & S_2{}^{(ij)}: & g_{i,j+2} = g_{i,j+1} g_{i,j} g^{-1}_{i,j+1}; \\
(9) & S^{(ij)}_{12}: & \alpha_{ij} g_{i+1,j} g_{i+1,j-1} \alpha^{-1}_{i,j-1} g_{i,j-1} g^{-1}_{ij} = 1.
\end{array}$$

The relations (8_2) imply that $g_{i,j+1} g_{i,j}$ is independent of j. Let for brevity,

$$(10) \qquad g_{i,j+1} g_{i,j} = s_i.$$

The recurrence relations (9) allow us to express all α_{ij}'s in terms of the $g_{\alpha\beta}$'s, $\alpha_{i,0}$. Taking into account (5a) and (10) we find

$$(9') \qquad \alpha_{ij} = g_{ij} g^{-1}_{i0} s^{-j}_{i+1}, \qquad (i, j = 0, \pm 1, \pm 2, \cdots).$$

Substituting these expressions of the α_{ij}'s into the relations (7) and taking into account (10) we find in a straightforward manner that the relations (7) can be replaced by the following relations:

$$(7') \quad g_{i1} g_{i+1,1} g_{i+1,0} g_{i,0} g_2 \cdots g_{n-2} g^2{}_{n-1} g_{n-2} \cdots g_2 = 1, \quad (i = 0, \pm 1, \pm 2, \cdots).$$

We have thus obtained a first reduction of the algebraic expression of the group H_n: *as generators of H_n we have the elements g_{ij}, $(i, j = 0, \pm 1, \pm 2, \cdots)$, $g_2, \cdots, g_{n-1}$; the generating relations are* (5b), (6)-(6'''), (7'), (8_1), (8_2), *where the elements α_{ij} in (8_1) are defined by* (9'), *and* (10).

Since $\alpha_{i0} = 1$, the relation (8_1) for $j = 0$ yields the following relation:

(11) $$g_{i+1,0}g_{i,0} = g_{i,0}g_{i-1,0}, \qquad (i = 0, \pm 1, \pm 2, \cdots).$$

Since, by (7′), the product $g_{i,1}g_{i+1,1}g_{i+1,0}g_{i,0}$ is independent of i, we deduce, as a consequence of (11), the following relation:

(11′) $$g_{i,1}g_{i+1,1} = g_{i-1,1}g_{i,1}, \qquad (i = 0, \pm 1, \pm 2, \cdots).$$

We proceed to prove that *in the reduced complete set of generating relations of H_n given above, the infinite set of relations* (8_1) *can be replaced by the relations* (11) *and* (11′), *i. e. the relations* (8_1) *are group-theoretic consequences of the remaining generating relations and of* (11) *and* (11′).

Proof. We denote by τ the product $g_2 \cdots g_{n-2}g^2{}_{n-1}g_{n-2} \cdots g_2$, or, in view of (7′),

(12) $$\tau = (g_{i1}g_{i+1,1}g_{i+1,0}g_{i,0})^{-1}.$$

We have

$$\tau^{-1} = g_{i1}s_{i+1}g_{i0} = s_i g_{i0}^{-1}s_{i+1}g_{i0},$$

hence

$$s_{i+1}^{-1} = g_{i,0}\tau s_i g_{i,0}^{-1}.$$

Substituting into (9′) we get

$$\alpha_{ij} = g_{ij}(\tau s_i)^j g_{i,0}^{-1},$$

and hence, using a new symbol β_{ij} for the transforming elements in (8_1), we have

(13) $$\beta_{ij} = \alpha_{ij}g_{i+1,j}^{-1}\alpha_{i+1,j} = g_{ij}(\tau s_i)^j g_{i0}^{-1}(\tau s_{i+1})^j g_{i+1,0}^{-1}.$$

By (5b) each element α_{ij} is commutative with τ, since $\tau = g_2 \cdots g^2{}_{n-1} \cdots g_2$, hence τ is also commutative with $\alpha_{ij}\alpha_{i,j-1}^{-1}$, i. e. in view of (7) (which is a consequence of (7′) and (9′)), τ is commutative with $g_{ij}\tau g_{ij}$:

(14) $$(\tau g_{ij})^2 = (g_{ij}\tau)^2.$$

We use the following relations:

(15) $$(\tau s_i)g_{i0}^{-1} = \tau g_{i1} \quad [\text{by } (10)]$$

(15′) $$\tau s_i(\tau g_{i1})(\tau s_{i+1}) = \tau g_{i2}\tau g_{i1}g_{i0}^{-1} \quad [\text{by } (10), (12) \text{ and } (14)]$$

(15″) $$\tau s_i(\tau g_{ij}\tau g_{i,j-1}) = (\tau g_{i,j+1}\tau g_{ij})\tau s_i \quad [\text{by } (10) \text{ and } (14)].$$

To prove (15′) and (15″) we proceed as follows:

$$s_i\tau g_{ij}\tau = g_{i,j+1}g_{ij}\tau g_{ij}\tau \quad [\text{by } (10)] = g_{i,j+1}\tau g_{ij}\tau g_{ij};$$

hence

$$\tau s_i(\tau g_{i1})(\tau s_{i+1}) = \tau g_{i2}\tau g_{i1}\tau g_{i1}s_{i+1} = \tau g_{i2}\tau g_{i1}g_{i0}^{-1} \quad [\text{by } (12)],$$

and

$$\tau s_i(\tau g_{ij}\tau g_{i,j-1}) = \tau g_{i,j+1}\tau g_{ij}\tau g_{ij}g_{i,j-1} = \tau g_{i,j+1}\tau g_{ij}\tau s_i.$$

From (15), (15′), and (15″) we deduce for any integer $k \geq 0$ the following relations:

$$\left.\begin{array}{ll}(16) & \gamma_{i,2k} = (\tau s_i)^{2k} g_{i,0}^{-1}(\tau s_{i+1})^k = \tau g_{i,2k} \cdots \tau g_{i1} g_{i0}^{-1}; \\ (16') & \gamma_{i,2k+1} = (\tau s_i)^{2k+1} g_{i,0}^{-1}(\tau s_{i+1})^k = \tau g_{i,2k+1} \cdots \tau g_{i,1}.\end{array}\right\} \quad k \geqq 0$$

In an exactly similar manner the following relations can be verified:

$$(\tau s_{i+1}) g_{i0} (\tau s_i) = \tau g_{i0};$$
$$\tau g_{i0} \cdot \tau s_i = g_{i0} \cdot \tau g_{i0} \tau g_{i,-1}.$$

From these relations and from (15″) we deduce for any integer $k < 0$ the following relations:

$$\left.\begin{array}{ll}(17) & \gamma_{i,2k} = (\tau s_i)^{2k} g_{i0}^{-1}(\tau s_{i+1})^k = g_{i,2k+1}^{-1}\tau^{-1} g_{i,2k+2}^{-1}\tau^{-1} \cdots g_{i,0}^{-1}\tau^{-1}g_{i0}^{-1}; \\ (17') & \gamma_{i,2k+1} = (\tau s_i)^{2k+1} g_{i0}^{-1}(\tau s_{i+1})^k = g_{i,2k+2}^{-1}\tau^{-1} \cdots g_{i0}^{-1}\tau^{-1}.\end{array}\right\} \quad k < 0.$$

Using the relations (16), (16′), (17) and (17′) and taking into account the relations (8_2) and (14), we obtain easily the following relations:

$$\left.\begin{array}{ll}(18) & \gamma_{i,2k}^{-1} g_{i,2k} \gamma_{i,2k} = g_{i0} \\ (18') & \gamma_{i,2k+1}^{-1} g_{i,2k+1} \gamma_{i,2k+1} = \tau g_{i1}\tau^{-1}\end{array}\right\} k \gtreqless 0.$$

We put

$$(\tau s_{i+1})^k g_{i+1,0}^{-1} = \delta_{i,k}$$

so that, by (13),

$$\beta_{i,2k} = g_{i,2k} \gamma_{i,2k} \delta_{i,k},$$
$$\beta_{i,2k+1} = g_{i,2k+1}\gamma_{i,2k+1}\delta_{i,k+1}.$$

In view of (18) and (18′), the relations (8_1) will follow at once, if we establish the following relations:

$$(19) \qquad \delta_{ik}^{-1} g_{i0} \delta_{ik} = g_{i+2,2k},$$
$$(19') \qquad \delta_{ik}^{-1} \tau g_{i1} \tau^{-1} \delta_{ik} = g_{i+2,2k-1}.$$

Since $\delta_{i0} = g_{i+1,0}^{-1}$ and $\tau^{-1}\delta_{i1} = s_{i+1} g_{i+0,0}^{-1} = g_{i+1,1}$, the relation (19) for $k = 0$ coincides with (11) and the relaticns (19′) for $k = 1$ coincides with (11′). Hence in order to establish the relations (19) and (19′) for all k it is sufficient to show that if they hold true for a given k, they also hold true for $k + 1$ and for $k - 1$. Now

$$\begin{aligned}\delta_{i,k+1} &= \delta_{ik} g_{i+1,0} \tau s_{i+1} g_{i+1,0}^{-1} = \delta_{ik} g_{i+1,0} (g_{i+1,1} g_{i+2,1} g_{i+2,0} g_{i+1,0})^{-1} g_{i+1,1} \\ &= \delta_{ik} (g_{i+2,1} g_{i+2,0})^{-1} = \delta_{ik} s_{i+2}^{-1}.\end{aligned}$$

Hence, assuming (19) and (19′) for a given value of k, the same relations follow also for $k+1$ and $k-1$ in view of the relations

$$s^{-\epsilon}_{i+2} g_{i+2,j} s^{\epsilon}_{i+2} = g_{i+2,j-2\epsilon}, \qquad \epsilon = \pm 1$$

which are direct consequences of the relations (8_2), q. e. d.

7. We now complete the elimination of the elements α_{ij} from the generating relations of H_n by proving that also the commutativity relations $\alpha_{ij} g_k = g_k \alpha_{ij}$, $k > 2$, (5b), are consequences of the remaining relations, to wit, of the relations (6), (6″), (8_2), and (7′). This is obvious if $k > 3$, since the α_{ij} depend only on the g_{ij}'s [see (9′)] and since, by (6), the g_{ij}'s are commutative with $g_3, \cdots, g_{n-1}$. By (7), which is a consequence of (10) and of the relations (9′) which define the elements α_{ij}, we have

$$\alpha_{ij}\alpha^{-1}_{i,j-1} = g_{ij} g_2 \cdots g^2_{n-1} \cdots g_2 g_{ij}.$$

Since $\alpha_{i0} = 1$, it is sufficient to establish the commutativity of $\alpha_{ij}\alpha^{-1}_{i,j-1}$ and g_2. Now, using the relations (6) and (6‴) we find:

$$\begin{aligned} g_2 \cdot g_{ij} g_2 g_3 \cdots g^2_{n-1} \cdots g_3 g_2 g_{ij} &= g_{ij} g_2 g_{ij} g_3 \cdots g^2_{n-1} \cdots g_3 g_2 g_{ij} \\ = g_{ij} g_2 \cdots g^2_{n-1} \cdots g_3 g_{ij} g_2 g_{ij} &= g_{ij} g_2 \cdots g^2_{n-1} \cdots g_2 g_{ij} g_2, \end{aligned}$$

and this proves our assertion.

Reassuming the reduction carried out so far, we have that our group H_n is defined by the following set of generators and generating relations:

Generators:

(20) $\qquad g_{ij}\ (i, j = 0, \pm 1, \pm 2, \cdots), \qquad g_2, \cdots, g_{n-1}.$

Generating relations:

(21) $\qquad g_{i,j+2} = g_{i,j+1} g_{ij} g^{-1}_{i,j+1}$

(22) $\qquad g_{i+2,0} = g_{i+1,0} g_{i,0} g^{-1}_{i+1,0}$

(22_1) $\qquad g_{i+2,1} = g^{-1}_{i+1,1} g_{i,1} g_{i+1,1}$

(23) $\qquad g_{ij} g_2 g_{ij} = g_2 g_{ij} g_2$

(23_1) $\qquad g_{ij} g_k = g_k g_{ij} \qquad (k = 3, \cdots, n-1)$

(24) $\qquad g_k g_{k+1} g_k = g_{k+1} g_k g_{k+1} \qquad (k = 2, 3, \cdots, n-2)$

(24_1) $\qquad g_k g_l = g_l g_k, \quad | k - l | \neq 1$

(25) $\qquad g_{i1} g_{i+1,1} g_{i+1,0} g_{i,0} g_2 \cdots g_{n-2} g^2_{n-1} g_{n-2} \cdots g_2 = 1.$

The existence of a finite set of generators follows now readily. In fact, the relations (21), for a fixed value of i, can be considered as recurrence relations which define the elements g_{ij} in terms of the two free elements g_{i0} and g_{i1}. Then the relations (22) and (22_1) can be used in order to express all the elements g_{i0} and g_{i1} in terms of g_{00}, g_{10} and g_{01}, g_{11}, respectively. *Consequently our group H_n is generated by the $n+2$ elements:*

$$(26) \qquad g_{00}, g_{10}, g_{01}, g_{11}, g_2, \cdots, g_{n-1}.$$

The reduction of the infinite set of relations (23_1) is trivial: since all the g_{ij}'s are expressible in terms of g_{00}, g_{10}, g_{01}, g_{11}, all the relation (23_1) are consequences of the four relations

$$(23_2) \qquad g_{ij}g_k = g_kg_{ij}, \qquad (i, j = 0, 1).$$

The reduction of the relations (23) is based on the following

LEMMA. *If four elements a, b, c, x* satisfy the relations

$$axa = xax, \quad bxb = xbx, \quad cxc = xcx, \quad c = b^{-\epsilon}ab^{\epsilon}, \quad \epsilon = 1 \text{ or } -1,$$

and if $d = c^{-\epsilon}bc^{\epsilon}$, then the above relations have as a group-theoretic consequence the relation $xdx = dxd$.

Proof. Let $\epsilon = +1$. Then

$$\begin{aligned} bcxc^{-1}b^{-1} &= bx^{-1}cxb^{-1} = bx^{-1}b^{-1}abxb^{-1} = x^{-1}b^{-1}xax^{-1}bx \\ &= x^{-1}b^{-1}a^{-1}xabx = x^{-1}c^{-1}b^{-1}xbcx. \end{aligned}$$

Hence

$$\begin{aligned} dxd^{-1} &= c^{-1}bcxc^{-1}b^{-1}c = x^{-1}c^{-1}x^{-1}b^{-1}xbxcx = x^{-1}c^{-1}x^{-1}xbx^{-1}xcx \\ &= x^{-1}c^{-1}bcx = x^{-1}dx. \end{aligned}$$

The case $\epsilon = -1$ is reducible to the case $\epsilon = +1$, by the substitution

$$x_1 = x^{-1}, \quad a_1 = a^{-1}, \quad b_1 = b^{-1}; \quad c_1 = c^{-1}.$$

Putting $a = g_{ij}$, $b = g_{i,j+1}$, $c = g_{i,j+2}$, $\epsilon = -1$, we deduce from the above lemma, in view of (21), that for a fixed i, the relations (23) relative to the indices j, $j+1$, $j+2$ imply as a group-theoretic consequence the relation (23) for the index $j+3$. Similarly, if we put $a = g_{i,j+2}$, $b = g_{i,j+1}$, $c = g_{ij}$, $\epsilon = 1$, we find that the just mentioned three consecutive relations (23) also imply the relation (23) for the index $j-1$. It follows, that for a fixed i, all the relations (23) are consequences of any three of them relative to three consecutive indices, say $j = 0, 1, 2$:

$$g_{ij}g_2g_{ij} = g_2g_{ij}g_2, \qquad (j = 0, 1, 2). \tag{27}$$

Now, in view of (22) and (22_1), we conclude in a similar manner, on the basis of the preceding lemma, that for $j = 0, 1$ the relations $g_{ij}g_2g_{ij} = g_2g_{ij}g_2$ are consequences of three of these relations relative to three consecutive valus of i, say $i = 0, 1, 2$. It remains to consider the set of relations $g_{i2}g_2g_{i2} = g_2g_{i2}g_2$. Using the relations (27) for $j = 0, 1$ and the expression of g_{i2} derived from (21), for $j = 0$, we change the above relation into an equivalent relation as follows:

$$\begin{aligned} g_{i2}g_2g_{i2}g_2^{-1}g_{i2}^{-1}g_2^{-1} &= g_{i1}g_{i0}\underline{g_{i1}^{-1}g_2g_{i1}}g_{i0}\underline{g_{i1}^{-1}g_2^{-1}g_{i1}}g_{i0}^{-1}g_{i1}^{-1}g_2^{-1} \\ &= g_{i1}g_{i0}g_2g_{i1}\underline{g_2^{-1}g_{i0}g_2}g_{i1}^{-1}g_2^{-1}g_{i0}^{-1}g_{i1}^{-1}g_2^{-1} = (g_{i1}g_{i0}g_2)^2(g_2g_{i1}g_{i0})^{-2}. \end{aligned}$$

Hence the relation $g_{i2}g_2g_{i2} = g_2g_{i2}g_2$ can be replaced by the relation

$$(g_{i1}g_{i0}g_2)^2 = (g_2g_{i1}g_{i0})^2. \tag{28}$$

Now, it is not difficult to see that the expressions $\sigma_i = (g_{i1}g_{i0}g_2)^2(g_2g_{i1}g_{i0})^{-2}$ are all transforms of each other, for $i = 0, \pm 1, \pm 2, \cdots$, as a consequence of the relations (27) ($j = 0, 1$), (23_1) and (25). In fact, let

$$g_3 \cdots g_{n-2}g^2_{n-1}g_{n-2} \cdots g_3 = \delta,$$

so that, by (23_1), we have $g_{ij}\delta = \delta g_{ij}$. By (25), we have

$$g_{i+1,1}g_{i+1,0} = (g_{i0}g_2\delta g_2g_{i1})^{-1}.$$

Hence, substituting into σ_{i+1} we find

$$\begin{aligned} \sigma_{i+1} &= (g_{i1}^{-1}g_2^{-1}\delta^{-1}g_2^{-1}g_{i0}^{-1}g_2)^2(g_{i0}g_2\delta g_2g_{i1}g_2^{-1})^2 \\ &= g_{i1}^{-1}g_2^{-1}\delta^{-1}g_2^{-1}g_{i0}^{-1}\underline{g_2g_{i1}^{-1}g_2^{-1}\delta^{-1}g_{i0}\delta}g_2g_{i1}\underline{g_2^{-1}g_{i0}g_2\delta g_2g_{i1}g_2^{-1}} \\ &= g_{i1}^{-1}g_2^{-1}\delta^{-1}g_2^{-1}(g_2g_{i1}g_{i0})^{-1}(g_{i1}g_{i0}g_2)^2g_{i0}^{-1}g_{i1}^{-1}\delta g_2g_{i1} \\ &= (g_{i0}^{-1}g_{i1}^{-1}\delta g_2g_{i1})^{-1}\bar{\sigma}_i(g_{i0}^{-1}g_{i1}^{-1}\delta g_2g_{i1}), \end{aligned}$$

where $\bar{\sigma}_i = (g_2g_{i1}g_{i0})^{-2}(g_{i1}g_{i0}g_2)^2$, obviously a transform of σ_i.

Hence, any one of the relations (28) implies as a consequence the entire set of these relations. We shall take the relations relative to $i = 0$.

We finally observe that in view of the relations $g_{i1}g_{i+1,1} = g_{i+1,1}g_{i+2,1}$ and $g_{i+1,0}g_{i,0} = g_{i+2,0}g_{i+1,0}$ [(22_1) and (22) respectively], the infinite set of relations (25) reduces to one relation, say relative to $i = 0$.

Reassuming, we have the following result:

The group H_n is defined by the following set of generators and generating relations:

1. *Generators:*

$$(29) \qquad g_{00}, g_{10}, g_{01}, g_{11}, g_2, \cdots, g_{n-1}$$

2. *Generating relations:*

$$(30) \qquad g_{ij}g_2g_{ij} = g_2g_{ij}g_2, \qquad (i, j = 0, 1).$$

$$(30') \qquad (g_{10}g_{00}g_2)^2 = (g_2g_{10}g_{00})^2$$

$$(30'') \qquad (g_{01}g_{11}g_2)^2 = (g_2g_{01}g_{11})^2$$

$$(30''') \qquad (g_{01}g_{00}g_2)^2 = (g_2g_{01}g_{00})^2$$

$$(31) \qquad g_{ij}g_k = g_kg_{ij}, \qquad k > 2$$

$$(32) \qquad g_kg_{k+1}g_k = g_{k+1}g_kg_{k+1}, \qquad (k = 2, 3, \cdots, n-2)$$

$$(32') \qquad g_kg_l = g_lg_k, \qquad |k - l| \neq 1$$

$$(33) \qquad g_{01}g_{11}g_{10}g_{00}g_2 \cdots g_{n-2}g^2_{n-1}g_{n-2} \cdots g_2 = 1.$$

The relations (30′), (30″) and (30‴) are the relations (27) relative to the following values of the indices i, j: $i = 2, j = 0$; $i = 2, j = 1$; $i = 0, j = 2$, after the expressions of g_{20}, g_{21}, g_{02}, given by (22), (22_1) and (21) respectively, are substituted.

Remark. If we change our notation as follows:

$$(34) \qquad g_{01} = \lambda_1, \quad g_{11} = \lambda_2, \quad g_{10} = \lambda_3, \quad g_{00} = \lambda_4,$$

then we see that the relations (30′), (30″), (30‴) are of the form

$$(35) \qquad (g_2\lambda_i\lambda_j)^2 = (\lambda_i\lambda_jg_2)^2, \quad i < j, \quad i, j = 1, 2, 3, 4.$$

An easy verification shows that *the six relations* (35) *all hold true.* For $i = 3, j = 4$; $i = 1, j = 2$; $i = 1, j = 4$, they coincide with the relations (30′), (30″), (30‴) respectively. For $i = 2, j = 3$ the relation (35) coincides with (28) for $i = 1$. The reader will easily verify the truth of the relations (35) in the remaining two cases.

III. The fundamental group of plane elliptic curves.

8. What is the geometric significance of the invariant subgroup H_n of G_n? We proceed to show that H_n is the fundamental group of a certain subspace of $R^n - D$ (see section 1).

Let us consider R as the Riemann surface of some algebraic elliptic curve f. This curve carries, for every n, a simple infinity of complete linear series g_n^{n-1} of dimension $n - 1$, and each set of n points of f belongs to one and only

one series g_n^{n-1}. The simple infinity of these series is an elliptic one-dimensional variety, birationally equivalent to f: in fact, each series contains a unique set of n points of which $n-1$ are preassigned fixed points $P_1^0, \cdots, P^0_{n-1}$, and then the n-th point P of the set determines the series uniquely.

Since a g_n^{n-1} is represented on R^n by a space homeomorphic to a linear space of $n-1$ dimensions, we conclude that *the algebraic variety R^n contains an elliptic pencil $\{S_{n-1}\}$ of linear $(n-1)$-spaces S_{n-1} free from base points.*

Let V_{n-2} be the intersection of an S_{n-1} with the discriminant variety D of R^n. We assert that *H_n is the fundamental group of the residual space $S_{n-1} - V_{n-2}$.*

Proof. We take the origin O of the fundamental group to be a point of $S_{n-1} - V_{n-2}$. We have to prove that: (1) a singular 1-sphere γ on O and in $R^n - D$ represents an element of H_n, if and only if it can be deformed over $R^n - D$ into a 1-sphere γ' contained in $S_{n-1} - V_{n-2}$, the point O being fixed; (2) if γ is already in $S_{n-1} - V_{n-2}$ and if it bounds a singular 2-cell on $R^n - D$, then it also bounds a singular 2-cell on $S_{n-1} - V_{n-2}$.

Let u be an elliptic integral of the first kind attached to the curve f. It is well known that f admits a continuous one-parameter group of birational transformations π_t into itself, represented analytically by the equation $u' \equiv u + t$ (mod. periods). Each transformation π_t of the group is an automorphism of R. There is an induced automorphism of R^n, which we shall also denote by π_t and which is at the same time an automorphism of the residual space $R^n - D$, since π_t transforms sets of n distinct points of R into sets of n distinct points. Each π_t permutes the linear spaces S_{n-1} (images of linear series on f) and induces a birational transformation σ_τ into itself (an automorphism) of the elliptic pencil $\{S_{n-1}\}$. If we put

$$v = u(x_1) + u(x_2) + \cdots + u(x_n),$$

where $x_1, \cdots, x_n$ is an n-tuple of points of R, then v is a simple integral attached to R^n and v reduces to a constant on each member of the pencil S_{n-1} (theorem of Abel). Hence v is also an elliptic integral of the first kind attached to the pencil $\{S_{n-1}\}$, and the transformation σ_τ is given by the equation

$$v' \equiv v + \tau, \qquad \tau \equiv nt.$$

The group of the transformations π_t covers n^2 times the group of transformations σ_τ, since to the identity σ_0 correspond the n^2 transformations $\pi_{\omega/n}$, where ω/n is the n-th of a period of u. Since the covering is free from branch points, it follows immediately that *any variation of S_{n-1} in the pencil $\{S_{n-1}\}$ can be*

accompanied by an isotopic deformation of the variable S_{n-1}, *and in such a manner that also the residual space is deformed isotopically.* This isotopic deformation of $S_{n-1} - V_{n-2}$ is simply effected by a convenient chain of transformations π_t, $0 \leq t \leq 1$, applied to the initial position of the S_{n-1}. From this last statement we derive immediately the following conclusion. Let R' be the Riemann surface of the elliptic pencil $\{S_{n-1}\}$. Every point P of R^n lies on a definite S_{n-1} of the pencil and is thus mapped upon a definite point P' of R'. Similarly any point set A on R^n is mapped upon a point set A' of R'. A deformation of A on R^n or on $R^n - D$ induces a deformation of A' on R'. From the preceding statement we may conclude that, conversely, *if* A *is on* $R^n - D$, *then any deformation of* A' *on* R' *is induced by a deformation of* A *on* $R^n - D$, *and that if a point* P *of* A' *is fixed throughout the deformation of* A', *then the points of* A *which are mapped upon* P *may also be assumed to be fixed during the deformation of* A.

Let γ be a singular 1-sphere on $R^n - D$ issued for the origin O of the group G_n. From the definition of the group H_n follows that γ represents an element of H_n if and only if the map γ' of γ on R' is a (singular) 1-cycle ~ 0. Assume that $\gamma' \sim 0$. Then γ' can be contracted on R' to the point O', image of O, and hence, by our preceding result, γ can be deformed into a 1-sphere γ_1 contained in $S_{n-1} - V_{n-2}$, the point O being fixed. Conversely, if γ can be deformed into such a 1-sphere γ_1, then γ' can be contracted to the point O', $\gamma' \sim 0$, and hence γ represents an element of H_n.

Assume that γ is in $S_{n-1} - V_{n-2}$ and that it bounds a (singular) 2-cell E_2 on $R^n - D$. The map of E_2 on R' is a 2-sphere M' containing O', since the boundary of E_2 is mapped on the point O'. On R' any 2-sphere on O' can be contracted to the point O', this last point being fixed. Hence E_2 can be deformed on $R^n - D$ into a 2-cell contained in $S_{n-1} - V_{n-2}$, the boundary γ being fixed. This completes the proof of our theorem.

9. The variety V_{n-2} is an hypersurface immersed in the $(n-1)$-space S_{n-1}. Let S_2 be a general plane of S_{n-1} and let C be the plane algebraic curve along which S_2 cuts V_{n-2}. By a theorem proved in Zariski,[5] the fundamental group H_n of the residual space $S_{n-1} - V_{n-2}$ coincides with the fundamental group of $S_2 - C$. Now, a general plane S_2 in our S_{n-1} is the image of a general series $g_n{}^2$ immersed in the corresponding $g_n{}^{n-1}$ of the elliptic curve. A point of S_2 represents a set of the series $g_n{}^2$. If we refer the sets of the $g_n{}^2$ to the lines of a plane, we obtain the general plane elliptic curve Γ, of order n, a birational transform of f, on which the sets of the $g_n{}^2$ are cut out by the lines of the plane. The points of S_2 which are on C represent n-tuples of the $g_n{}^2$

with coincident points, hence correspond to the tangent lines of Γ. We conclude that *the plane curve C is the dual of a general elliptic plane curve Γ of order n, and that our group H_n is the fundamental group of the residual space of C.*

The curve C is of order $2n$, possesses $k = 3n$ cusps and $d = 2n(n-3)$ nodes, and is the maximal cuspidal elliptic curve of its order. Of the generating relations of H_n, the relations (30), (30′), (30″), (30‴), and (32) are all of the form $aba = bab$ and arise from the cusps of C. The commutativity relations (31) and (32′) are due to the nodes of C. Finally, the relations (33) corresponds to the relation $\lambda_1\lambda_2 \cdots \lambda_{2n} = 1$, where the λ_i's are loops contained in a general line of the plane of C and surrounding the $2n$ intersections of this line with C (see Zariski [2]). The relations $\lambda_5 = \lambda_{2n}$, $\lambda_6 = \lambda_{2n-1}$ etc., arise from the n tangent lines of C belonging to a pencil of lines (compare the analogous discussion of the singularities and of the corresponding generating relations in Zariski,[4] p. 615). By watching the effect which the removal of a cusp or of a node has upon the fundamental group, we arrive at conclusions relative to the fundamental group of any plane curve C' which admits C as a limiting case, in particular of any plane elliptic curve of even order $2n$, possessing only nodes and cusps (compare Zariski,[4] p. 616). Let us first remove a node, i. e. let us consider a node of C as virtually nonexistent. For the fundamental group this amounts to replacing a commutativity relation $ab = ba$ by the relation $a = b$. We may assume that the relation thus affected is the relation $g_2g_4 = g_4g_2$. We have then $g_4 = g_2$. Since $g_2g_5 = g_5g_2$ and $g_4g_5g_4 = g_5g_4g_5$, it follows $g_4 = g_5$. In a similar manner we find $g_2 = g_3 = g_4 = \cdots = g_{n-1}$. Since $g_{ij}g_2g_{ij} = g_2g_{ij}g_2$ and $g_{ij}g_3 = g_3g_{ij}$, the relation $g_2 = g_3$ implies the relation $g_{ij} = g_2$. Hence the fundamental group becomes a cyclic group of order $2n$. Let us now remove a cusp, by converting the cusp into a node. A relation $aba = bab$ will be affected and will have to be replaced by $a = b$. We may assume that the affected relation is the relation $g_{00}g_2g_{00} = g_2g_{00}g_2$. We have then $g_2 = g_{00}$, after the cusp has been removed. If $n \geq 4$, we may use the relations $g_{00}g_3 = g_3g_{00}$, $g_2g_3g_2 = g_3g_2g_3$ and we then find that $g_2 = g_3$. We conclude as before that the group becomes cyclic. Hence, we have the following result: *If C' is a plane curve of order $2n > 6$ with nodes and cusps, and if C' admits the maximal cuspidal elliptic curve C, of the same order, as a limiting case, without being a curve C itself, then the fundamental group of C' is cyclic* (*of order $2n$*). *In particular, every plane elliptic curve of even order $2n$ possessing less than $3n$ cusps has a cyclic fundamental group.*

In the exceptional case $n = 3$ we are dealing with the dual of a general

plane cubic, i. e. with an elliptic sextic having 9 cusps. We write the generating relations of H_3, using (27) instead of the equivalent set of relations (30)-(30‴):

(34) $$g_{ij}g_2g_{ij} = g_2g_{ij}g_2, \qquad (i, j = 0, 1, 2).$$

(35) $$g_{01}g_{11}g_{10}g_{00}g_2{}^2 = 1,$$

where

(36) $$\begin{cases} g_{i2} = g_{i1}g_{i0}g_{i1}^{-1} \\ g_{20} = g_{10}g_{00}g_{10}^{-1} \\ g_{21} = g_{11}^{-1}g_{01}g_{11}. \end{cases}$$

The 9 relations (34) are typical cuspidal relations, and one may conjecture that they correspond to the 9 cusps of the curve. However, since only 7 of these relations are group-theoretically independent, this conjecture requires proof. The 7 independent relations are given by (30), (30′), (30″), (30‴) and correspond to the following values of i, j:

$$i, j = 0, 1; \quad i = 2, j = 0; \quad i = 2, j = 1; \quad i = 0, j = 2.$$

At present we can only assert that the seven independent relations are relations at 7 of the cusps. We recall that the well known group of 9 flexes of a cubic curve is doubly transitive. Hence if we remove a certain number of cusps of C, it is immaterial which cusps are removed, as long as the number of removed cusps does not exceed 2.

We remove successively the two cusps which give rise to the relations $g_{00}g_2g_{00} = g_2g_{00}g_2$, $g_{10}g_2g_{10} = g_2g_{10}g_2$. After the removal of the first cusp we have $g_{00} = g_2$. The relations (30′) and (30‴) become then consequences of the relations (30) for $i = 1$, $j = 0$ and $i = 0$, $j = 1$ respectively, while the relation (30″) becomes a consequence of (33). Hence the fundamental group of a sextic with 8 cusps (and with one or no double points) is generated by 4 elements

$$g_{10}, g_{01}, g_{11}, g_2$$

satisfying the relations:

$$g_{ij}g_2g_{ij} = g_2g_{ij}g_2$$
$$g_{01}g_{11}g_{10}g_2{}^3 = 1.$$

If we now remove the second cusp, we get $g_{10} = g_2$, and thus the fundamental group of a sextic with 7 cusps and of genus ≥ 1 is generated by 3 elements

$$g_{01}, g_{11}, g_2$$

satisfying the 3 relations:

$$g_{01}g_2g_{01} = g_2g_{01}g_2$$
$$g_{11}g_2g_{11} = g_2g_{11}g_2$$
$$g_{01}g_{11}g_2^4 = 1.$$

This is also a group generated by the two elements

$$u = g_{11}g_2g_{11}, \qquad v = g_{11}g_2$$

satisfying the relations $u^2 = v^3 = 1$.

It is known that this group is also the fundamental group of a sextic with six cusps on a conic (Zariski [2]). Hence there must be among the seven cusps left, a third cusp whose removal has no effect on the fundamental group. It is easily seen, by using the relations (36), that if the removal of an additional cusp yield the relation $g_{ij} = g_2$, $i \neq 2$, $j \neq 0$, the group becomes cyclic. On the contrary, if the removed cusp gave rise originally to the relation $g_{20}g_2g_{20} = g_2g_{20}g_2$, the group is unaltered, since the relation $g_{20} = g_2$ is already implied, in view of (36), by the removal of the first two cusps ($g_{00} = g_{10} = g_2$).

It is known that the sextics with six cusps distribute themselves into two distinct continuous systems, according as the six cusps lie or do not lie on a conic. The preceding considerations lead therefore to the conclusion that *the fundamental group of a sextic with six cusps not on a conic is cyclic* (of period 6).

One can verify the following: if $g_{i_1j_1}, g_{i_2j_2}, g_{i_3j_3}$ are any 3 of our nine elements g_{ij} such that $i_1 + i_2 + i_3 \equiv j_1 + j_2 + j_3 \equiv 0(3)$, then the removal of the three corresponding cusps (whence the addition of the relations $g_{i_1j_1} = g_{i_2j_2} = g_{i_3j_3} = g_2$) leads to a curve whose fundamental group is the above mentioned group of a sextic with six cusps on a conic. If, however, the above congruences do not hold true simultaneously, then the removal of the corresponding cusps leads to a curve with a cyclic fundamental group. What we have here is obviously something which adds topological significance to the configuration of the 12 MacLaurin lines determined by the nine flexes of a cubic curve. It is known that if the nine flexes are distributed into three triples lying on three MacLaurin lines, then the six flex tangents of any two of the triples lie on a line conic. Dually, any two of the corresponding triples of cusps lie on a conic. If then the three cusps of one triple are considered as virtual non-existent, the resulting sextic must have six cusps on a conic.

This proves incidentally, that the nine relations (34) reproduce exactly the relations at the nine cusps.

10. We conclude by pointing out that the reasoning employed in our paper,[4] section 7, can be applied also in the present case to elliptic curves of odd order and leads to the conclusion that the fundamental group of such a curve (with nodes and cusps) is always cyclic. For the proof it is sufficient to consider the maximal cuspidal elliptic curve C_{2n+1}, of order $2n+1$, and to observe that C_{2n+1} can be degenerated into the maximal cuspidal elliptic curve C_{2n} and into a line p tangent to C_{2n}. The fundamental group of $C_{2n}+p$ can be obtained from the fundamental group H_n of C_{2n} by adding an extra generator γ and the relations $(\gamma g_2)^2=(g_2\gamma)^2$, $\gamma g_{ij}=g_{ij}\gamma$, $\gamma g_k=g_k\gamma$, $k>2$. We obtain the curve C_{2n+1} by considering the tacnode of $C_{2n}+p$ at the point of tangency of p, as a virtual cusp. As a consequence, we replace the relation $(\gamma g_2)^2=(g_2\gamma)^2$ by the relation $g_2=\gamma$, and from this follows immediately that the group of C_{2n+1} is cyclic.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] K. Reidemeister, "Knoten und Gruppen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 5 (1927).

[2] O. Zariski, "On the problem of existence of algebraic functions of two variables possessing a given branch curve," *American Journal of Mathematics*, vol. 51 (1929).

[3] O. Zariski, "A topological proof of the Riemann-Roch theorem on an algebraic curve," *American Journal of Mathematics*, vol. 58 (1936).

[4] O. Zariski, "On the Poincaré group of rational plane curves," *American Journal of Mathematics*, vol. 58 (1936).

[5] O. Zariski, "A theorem on the Poincaré group of an algebraic hypersurface," *Annals of Mathematics*, vol. 38 (1937).

ON THOSE POINTS OF AN ALGEBRAIC MANIFOLD NOT REACHABLE BY A GIVEN PARAMETRIC REPRESENTATION.[1]

By J. F. Daly.

Let K denote the complex field, and let $x_1, \cdots, x_n$ be elements of an algebraic extension $K(t_1, \cdots, t_m, x_1, \cdots, x_n)$ of the pure transcendental field $K(t_1, \cdots, t_m)$. A set of complex numbers $t'_1, \cdots, t'_m, x'_1, \cdots, x'_n$ will be called allowable if every polynomial $F(t, x)$ which vanishes as an element of $K(t_1, \cdots, t_m, x_1, \cdots, x_n)$ vanishes also when t'_i, x'_i are substituted for t_i, x_i respectively. The totality of points whose coördinates $\{x'_1, \cdots, x'_n\}$ belong to allowable sets will be contained in some smallest algebraic manifold $\mathfrak{M}$. This manifold is said to be represented parametrically in terms of the t's.

In general $\mathfrak{M}$ will contain points which are not allowable. But in case the parameters are merely some r of the coördinates themselves it follows readily from a theorem of Ritt (1) that such exceptional points are always limit points of allowable points. It is the purpose of the present paper to extend the above result to all representations, whether the parameters t are essential or not.

Theorem. *If an algebraic manifold $\mathfrak{M}$ is represented in terms of any parameters whatever, the base field K being the complex field, then each point of $\mathfrak{M}$ is a limit of allowable points.*

We shall treat in detail only the case in which the parameters $t_1, \cdots, t_m$ are all algebraically independent over K; but the method of proof is quite the same if the representation involves additional parameters, say $u_{m+1}, \cdots, u_s$, dependent on the t's. Since each x is algebraic over any extension of $K(t_1, \cdots, t_m)$ we may write the irreducible equation for x_{a_1} over $K(t_1, \cdots, t_m)$, the irreducible equation for x_{a_2} over $K(t_1, \cdots, t_m, x_{a_1})$, etc., each divided through by its leading coefficient:

$$(\alpha)\qquad \begin{aligned}
&x^a_{a_1} + a_1(t_1, \cdots, t_m)x^{a-1}_{a_1} + \cdots + a_a(t_1, \cdots, t_m) = 0\\
&x^b_{a_2} + b_1(t_1, \cdots, t_m, x_{a_1})x^{b-1}_{a_2} + \cdots + b_b(t_1, \cdots, t_m, x_{a_1}) = 0\\
&\quad\vdots\\
&x^c_{a_n} + c_1(t_1, \cdots, t_m, x_{a_1}, \cdots, x_{a_{n-1}})x^{c-1}_{a_n} + \cdots + c_c(t_1, \cdots, x_{a_{n-1}}) = 0
\end{aligned}$$

[1] Received November 12, 1936.

the order $x_{a_1}, \cdots, x_{a_n}$ to be determined later. Although the coefficients $a_i, \cdots, c_i$ are in general rational functions of both x's and t's, they may be made polynomials in the x's; the denominators will then involve only $t_1, \cdots, t_m$. Let a non-vanishing common multiple of all denominators be $A(t_1, \cdots, t_m)$. Any set of complex numbers $t'_1, \cdots, t'_m, x'_1, \cdots, x'_n$ satisfying (α) together with the relation $A(t'_1, \cdots, t'_m) \neq 0$ is allowable (2).

The theorem will be proved by showing that for any point $\{x'_1, \cdots, x'_n\}$ of $\mathfrak{M}$ there is a neighboring point $\{\dot{x}_1, \cdots, \dot{x}_n\}$ of to which we can assign parameter values $\dot{t}_1, \cdots, \dot{t}_m$ in such a way that the set $\dot{t}_1, \cdots, \dot{t}_m, \dot{x}_1, \cdots, \dot{x}_n$ satisfies (α) with $A(\dot{t}_1, \cdots, \dot{t}_m) \neq 0$.

For this purpose we choose a new transcendental basis (3) of $K(t_1, \cdots, t_m, x_1, \cdots, x_n)$. Every element of $K(t_1 \cdots, t_m, x_1, \cdots, x_n)$ is algebraically dependent on the ordered set $x_1, \cdots, x_n, t_1, \cdots, t_m$. If from this set we select those elements which are algebraically independent (over K) of all preceding elements, we obtain a set Σ having the following properties:

(a) the number of elements in Σ is exactly m;

(b) every element of $K(t_1, \cdots, t_m, x_1, \cdots, x_n)$ is algebraically dependent on Σ;

(c) the elements of Σ are algebraically independent over K.

Let the elements of Σ be $x_{\beta_1}, \cdots, x_{\beta_r}, t_{\gamma_1}, \cdots, t_{\gamma_{m-r}}$. After suitably renumbering the x's and t's, we may assume that Σ contains $x_1, \cdots, x_r, t_1, \cdots, t_{m-r}$, and that equations (α) have been calculated correspondingly.

The field $K(t_1, \cdots, t_m, x_1, \cdots, x_n)$ can now be regarded as an algebraic extension of the pure transcendental field $K(x_1, \cdots, x_r, t_1, \cdots, t_{m-r})$. We may therefore write the irreducible equation satisfied by x_{r+1} over $K(x_1, \cdots, x_r, t_1, \cdots, t_{m-r})$, the irreducible equation satisfied by x_{r+2} over $K(x_1, \cdots, x_r, t_1, \cdots, t_{m-r}, x_{r+1})$, etc., each divided through by its leading coefficient. Note however that no t's will appear in the coefficients of the equations for $x_{r+1}, \cdots, x_n$; for the existence of a relation $F(x_1, \cdots, x_n, t_1, \cdots, t_{m-r}) = 0$ which actually involved some $t_k \in \Sigma$ would imply the dependence of that t_k on the set $x_1, \cdots, x_n, t_1, \cdots, t_{k-1}, t_{k+1}, \cdots, t_{m-r}$ and therefore on the set $x_1, \cdots, x_r, t_1, \cdots, t_{k-1}, t_{k+1}, \cdots, t_{m-r}$, which is impossible. The equations under consideration then take the form:

$$(\beta) \quad \begin{array}{l} x^d_{r+1} + d_1(x_1, \cdots, x_r) x^{d-1}_{r+1} + \cdots + d_d(x_1, \cdots, x_r) = 0 \\ \quad \cdot \qquad\qquad \cdot \qquad\qquad\qquad\qquad \cdot \qquad \cdot \\ \quad \cdot \qquad\qquad \cdot \qquad\qquad\qquad\qquad \cdot \qquad \cdot \\ \quad \cdot \qquad\qquad \cdot \qquad\qquad\qquad\qquad \cdot \qquad \cdot \\ x_n{}^e + e_1(x_1, \cdots, x_r, x_{r+1}, \cdots, x_{n-1}) x_n{}^{e-1} + \cdots + e_e(x_1, \cdots, x_{n-1}) = 0. \end{array}$$

The denominators of the various coefficients need involve only $x_1, \cdots, x_r$; let a non-vanishing common multiple of all denominators be $B(x_1, \cdots, x_r)$.

Continuing, we write the irreducible equation satisfied by t_{m-r+1} over $K(x_1, \cdots, x_r, t_1, \cdots, t_{m-r})$, the irreducible equation satisfied by t_{m-r+2} over $K(x_1, \cdots, x_r, t_1, \cdots, t_{m-r}, t_{m-r+1})$, etc., each divided through by its leading coefficient:

$$(\gamma)\quad \begin{array}{l} t^g_{m-r+1} + g_1(x_i, t_1, \cdots, t_{m-r})\, t^{g-1}_{m-r+1} + \cdots + g_g(x_i, t_1, \cdots, t_{m-r}) = 0 \\ \cdot \qquad\qquad \cdot \qquad\qquad \cdot \qquad\qquad \cdot \\ \cdot \qquad\qquad \cdot \qquad\qquad \cdot \qquad\qquad \cdot \\ \cdot \qquad\qquad \cdot \qquad\qquad \cdot \qquad\qquad \cdot \\ t_m{}^h + h_1(x_i, t_1, \cdots, t_{m-r}, t_{m-r+1}, \cdots, t_{m-1})\, t_m{}^{h-1} + \cdots + h_h(x, t_1, \cdots, t_{m-1}) = 0. \end{array}$$

Only $x_1, \cdots, x_r, t_1, \cdots, t_{m-r}$ need occur in the denominators of the coefficients. Let a non-vanishing common multiple of all denominators be $C(x_1, \cdots, x_r, t_1, \cdots, t_{m-r})$.

Any set $x'_1, \cdots, x'_n, t'_1, \cdots, t'_m$ of complex numbers satisfying (β) and (γ) with $B \cdot C \neq 0$ will be an allowable set, and will therefore satisfy the equations resulting from (α) on multiplication of each of the latter by $A(t_1, \cdots, t_m)$. It remains then, to ensure the non-vanishing of $A(t'_1, \cdots, t'_m)$. Now as an element of $K(t_1, \cdots, t_m, x_1, \cdots, x_n)$, A satisfies some irreducible equation over $K(x_1, \cdots, x_r, t_1, \cdots, t_{m-r})$:

$$A^s + \frac{p_1(x_1, \cdots, x_r, t_1, \cdots, t_{m-r})}{q_1(x_1, \cdots, x_r, t_1, \cdots, t_{m-r})} A^{s-1} + \cdots + \frac{p_s(x, t)}{q_s(x, t)} = 0.$$

Let $D(x_1, \cdots, x_r, t_1, \cdots, t_{m-r})$ be a non-vanishing common multiple of all denominators and of the numerator p_s of the last coefficient. Then $D(x'_1, \cdots, x'_r, t'_1, \cdots, t'_{m-r}) \neq 0$ implies $A(t'_1, \cdots, t'_m) \neq 0$ for any allowable set $x'_1, \cdots, x'_r, t'_1, \cdots, t'_m$. Thus the non-vanishing of the polynomial $B \cdot C \cdot D \equiv V(x_1, \cdots, x_r, t_1, \cdots, t_{m-r})$ implies the non-vanishing of all denominators so far considered.

We may arrange V according to power-products of the t's, and take a non-vanishing common multiple $P(x_1, \cdots, x_r)$ of the resulting coefficients. Now P does not vanish everywhere on the irreducible (2) manifold $\mathfrak{M}$, since otherwise it would vanish identically as an element of $K(t_1, \cdots, t_m, x_1, \cdots, x_n)$. Let $\{x'_1, \cdots, x'_n\}$ be any point of $\mathfrak{M}$. If $P(x'_1, \cdots, x'_n) = 0$, then by Ritt's theorem that point is a limit point of points $\{\dot{x}_1, \cdots, \dot{x}_n\}$ of $\mathfrak{M}$ for which $P \neq 0$. Suppose therefore that $P(x'_1, \cdots, x'_n) \neq 0$. Now equations (β) after multiplication by $B(x_1, \cdots, x_r)$ are satisfied by all allowable values of the x's, and thus constitute part of the equations defining the manifold $\mathfrak{M}$. But at the point under consideration $B \neq 0$, so that its coördinates actually satisfy equations (β).

The independent quantities $t'_1, \cdots, t'_{m-r}$ may be chosen in such a way that $V(x'_1, \cdots, x'_r, t_1, \cdots, t_{m-r}) \neq 0$. Using these values $x'_1, \cdots, x'_n$, $t'_1, \cdots, t'_{m-r}$, we may calculate successively the remaining t's from (γ). The set $x'_1, \cdots, x'_n, t'_1, \cdots, t'_m$, being allowable, will satisfy equations (α), for by construction $A(t'_1, \cdots, t'_m) \neq 0$.

Thus any point of $\mathfrak{M}$ at which $P \neq 0$ may be obtained directly from the original parametric representation; and any other point of $\mathfrak{M}$ is a limit point of points thus reachable.

PRINCETON UNIVERSITY.

REFERENCES.

(1) J. F. Ritt, "Differential equations," *American Mathematical Society Colloquium Publications*, vol. 14 (1932), p. 91. Also B. L. van der Waerden, "Zur algebraischen Geometrie III," *Mathematische Annalen*, vol. 108 (1933), pp. 694-698.

(2) For notation and definitions cf. B. L. van der Waerden, *Moderne Algebra*, vol. II, pp. 51-61; or B. L. van der Waerden, "Zur Nullstellentheorie der Polynomideale," *Mathematische Annalen*, vol. 96 (1926-27), pp. 183-208.

(3) B. L. van der Waerden, *Moderne Algebra*, vol. I, pp. 204-206.

A REMARK CONCERNING THE PARAMETRIC REPRESENTATION OF AN ALGEBRAIC VARIETY.[1]

By Oscar Zariski.

In his paper "On those points of an algebraic manifold not reachable by a given parametric representation," published in the present issue of this Journal, Mr. J. F. Daly treats the case in which the number of parameters in a given parametric representation of an algebraic variety exceeds the dimension of the variety; nor need the parameters belong to the field of algebraic functions defined by the variety. While this type of parametric representations is more general than the one treated heretofore explicitly in the literature (see, especially, van der Waerden, "Über irreduzible algebraische Mannigfaltigkeiten," *Mathematische Annalen*, vol. 108 (1933)), it may be pointed out that the generalization given by Daly can also be obtained by using some simple properties of rational transformations of varieties. The following proof is taken from the mimeographed notes of the algebraic geometry seminar conducted by Professor Lefschetz and myself in Princeton, 1934.

1. Let V be an irreducible algebraic r-dimensional variety in $S_m(y_1, \cdots, y_m)$ and let

$$x_k = R_k(y_i) = P_k(y_i)/Q(y_i), \qquad (k = 1, 2, \cdots, n) \tag{1}$$

be the equations of a rational transformation of V into an algebraic variety W in $S_n(x_1, \cdots, x_n)$. We assume, of course, that $Q \neq 0$ on V. The coördinates $\eta_1, \eta_2, \cdots, \eta_m$ of a generic point of V are elements of a field $\Omega = K(\eta_1, \cdots, \eta_m)$ of algebraic functions of r independent variables, where K is the field of complex numbers. The coördinates of a generic point of W are $\xi_k = R_k(\eta_i)$ and define a field $\Omega' = K(\xi_k)$, $\Omega' \leqq \Omega$. If ρ $(\leqq r)$ is the degree of transcendentality of Ω', then W is of dimension ρ.

Let $g(y_1, \cdots, y_m)$ be some polynomial which does not vanish identically on V, and let T denote the set of points of W which can be obtained directly from the equations (1) and which correspond to points of V at which $g \neq 0$, i.e. points of W which correspond to points of V at which $Q \neq 0$ and $g \neq 0$. We prove that *W is the closure of T.*

From the theory of fields it follows that $\Omega = \Omega'(t_1, \cdots, t_s, \eta)$, $s + \rho = r$,

[1] Received November 16, 1936.

where the t_j's are algebraically independent over Ω' and where η satisfies an algebraic equation $f(\xi_k, t_l, \eta) = 0$ with coefficients in K. Let

$$\eta_i = S_i(\xi_k, t_l, \eta)/M(\xi_k, t_l),$$

where S_i and M are polynomials, and let

$$N(Q) = L(\xi_k, t_l)/M(\xi_k, t_l), \qquad N(g) = G(\xi_k, t_l)/M(\xi_k, t_l)$$

be the norms over $K(\xi_k, t_l)$ of $Q(\eta_1, \cdots, \eta_m)$ and of $g(\eta_1, \cdots, \eta_m)$ respectively. Here $M(\xi_k, t_l) \neq 0$, and also $L(\xi_k, t_l) \neq 0$, $G(\xi_k, t_l) \neq 0$, since $Q(\eta_i) \neq 0$ and $g(\eta_i) \neq 0$, by hypothesis. Let then x^0 be a point of W at which the polynomials in $t: M(x^0{}_k, t_l)$, $L(x^0{}_k, t_l)$, $G(x^0{}_k, t_l)$ do not vanish identically. If $(t^0{}_1, \cdots, t^0{}_s)$ is any set of values of the t's at which these polynomials do not vanish, and if η^0 is a root of $f(x^0{}_k, t^0{}_l, \eta) = 0$, then $y^0{}_i = S_i(x^0{}_k, t^0{}_l, \eta^0)/M(x^0{}_k, t^0{}_l)$ are the coördinates of a point of V at which $Q \neq 0$, $g \neq 0$, and moreover $x^0{}_k = R_k(y^0{}_i)$. That is, any point (x) of W at which none of the polynomials $M(x, t)$, $L(x, t)$, $G(x, t)$ vanishes identically belongs to the set T. It follows then by a theorem of Ritt that every point of W is a limit point of points in T, q. e. d.

2. Let the coördinates $x_1, \cdots, x_n$ of a generic point of an algebraic ρ-dimensional variety W be algebraic functions of r parameters $t_1, \cdots, t_r$, independent over K. The variety W is a rational transform of the r-dimensional variety whose generic point is $(x_1, x_2, \cdots, x_n, t_1, \cdots, t_r)$. We identify this variety with the variety V of the preceding section. For the variety V the parameters t_l are merely some of the coördinates, and the points of V which cannot be reached by this parametric representation satisfy a certain equation $g = 0$, where $g = g(x, t)$ is a polynomial not identically zero on V. By the preceding section it follows that the points of W which cannot be reached by the given parametric representation are limit points of reachable points.

THE JOHNS HOPKINS UNIVERSITY.

ON THE CONSTRUCTION OF SYMMETRIC RULED SURFACES.[1]

By Arnold Emch.

1. *Introduction.* Let $S_3(x)$ denote a projective space of three dimensions with the homogeneous variables x_1, x_2, x_3, x_4 and $\phi_1, \phi_2, \phi_3, \phi_4$ the four elementary symmetric functions on these variables, then

$$\Sigma a_k \phi_1^{n_1} \phi_2^{n_2} \phi_3^{n_3} \phi_4^{n_4} = 0, \tag{1}$$

with the n_i as positive integers, and $n_1 + 2n_2 + 3n_3 + 4n_4 = n$ satisfied represents, by definition, a symmetric surface. Noted examples are the Cayley cubic $\Sigma 1/x_i = 0$ and Clebsch's diagonal surface. The writer has investigated surfaces of this kind in a number of papers. Two symmetric surfaces intersect obviously in a symmetric space curve, which is left invariant by the symmetric group G_{24} of collineations, since the surfaces producing it are invariant. In this connection I mention the discovery of a remarkable sextic of genus four which lies on 10 cubic cones.[2]

It is naturally interesting to enquire about the possibility of symmetric ruled surfaces. It is clear that such surfaces exist, because $\phi_1^2 + \lambda\phi_2 = 0$ represents a pencil of symmetric quadrics (admitting imaginary rulings). Then, of course, there is the special class of symmetric cones. If we admit the existence of general symmetric ruled surfaces, then by G_{24} a generic generatrix determines immediately 23 more on the surface. The first question then is what is the lowest order of surface on 24 lines belonging to the group G_{24}. Choose a generic line l, determined parametrically by

$$\rho x_i = a_i + \lambda b_i \qquad (i = 1, 2, 3, 4), \tag{2}$$

with (a) and (b) as two arbitrary distinct points. Substituting (2) in (1), l will lie on (1) when (2) is identically equal to zero for all values of λ. This leads to an equation of degree n in λ, which is satisfied for all values of λ when its $n+1$ coefficients vanish. Hence the number of coefficients in (1) must have a number of coefficients (effective) $\geqq n+1$. The experimental solution of the Diophantine equation gives for the orders $n = 1, 2, 3, 4, 5, 6, 7, \cdots$ the number of effective constants of (1): $0, 1, 2, 3, 4, 7, 9, \cdots$. This shows

[1] Received October 21, 1936.

[2] "Über eine besondere Raumkurve sechster Ordnung," *Monatshefte für Mathematik und Physik*, vol. 40 (1933), pp. 193-200.

that the first case in which the inequality is satisfied is for $n = 6$. There is still one constant available. Hence

THEOREM 1. *On a generic line l in S_3 there is a pencil of symmetric sextic surfaces. There are no such surfaces of lower order on l.*

For general ruled surfaces one must look for higher than the 6th order. To solve this problem we use the parametric representation.

2. *Construction of ruled symmetric surfaces.* Consider the equations

$$
(3)\qquad
\begin{aligned}
\rho x_1 &= \phi^*_1(\lambda_2, \lambda_3, \lambda_4) + \mu\psi^*_1(\lambda_2, \lambda_3, \lambda_4),\\
\rho x_2 &= \phi^*_2(\lambda_1, \lambda_3, \lambda_4) + \mu\psi^*_2(\lambda_1, \lambda_3, \lambda_4),\\
\rho x_3 &= \phi^*_3(\lambda_1, \lambda_2, \lambda_4) + \mu\psi^*_3(\lambda_1, \lambda_2, \lambda_4),\\
\rho x_4 &= \phi^*_4(\lambda_1, \lambda_2, \lambda_3) + \mu\psi^*_4(\lambda_1, \lambda_2, \lambda_3),
\end{aligned}
$$

in which $\phi^*(\alpha, \beta, \gamma)$ and $\psi^*(\alpha, \beta, \gamma)$ are symmetric polynomials in α, β, γ, from which the ϕ^*_i and ψ^*_i are obtained by replacing α, β, γ by $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ as indicated. Then the ϕ^*_i and ψ^*_i, except as to a possible change of signs throughout, permute in the same way as any chosen permutations of the λ's. Hence a permutation

$$
\begin{pmatrix} \lambda_1\, \lambda_2\, \lambda_3\, \lambda_4 \\ \lambda_i\, \lambda_k\, \lambda_l\, \lambda_g \end{pmatrix} \text{ induces the permutation } \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ \rho x_i & \rho x_k & \rho x_l & \rho x_g \end{pmatrix},
$$

in which $\rho = \pm 1$. For a definite set of values for the λ's the ϕ^*'s and ψ^*'s in (3) represent two points in S_3, and with μ variable (3) represents a straight line l. Thus

THEOREM 2. *On the application of G_{24} to a definite set $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ and with μ variable* (3) *represents 24 lines l_i. Each of these is met by six others of the same set.*

To prove the second part of the theorem, notice that a line l_i cuts the six planes $x_i - x_k = 0$ in six points P_{ik}. By the substitution (ik) this P_{ik} is not changed, but l_i is transformed into that of set of 24 which corrresponds to (ik).

To (3) we now adjoin

$$
(4)\qquad
\begin{aligned}
M(\lambda_1, \lambda_2, \lambda_3, \lambda_4) &= 0,\\
N(\lambda_1, \lambda_2, \lambda_3, \lambda_4) &= 0,
\end{aligned}
$$

two equations of degree m and n and symmetric in the λ's, and also homogeneous in order to make them of dimension 3. Then in a $S_3(\lambda)$-space (4)

represents two surfaces which we assume to intersect in a complete irreducible curve C_{mn}. To every point (λ) of this curve correspond by the G_{24} 24 points (including the point itself) of the same curve and by (3) 24 lines l_i. As (λ) describes C_{mn}, the l_i's generate a ruled surface R whose genus is the same as that of C_{mn}, because there exists a $(1,1)$ correspondence between the points of C_{mn} and the generatrices of the ruled surface in $S_3(x)$. To determine the order of R, the λ's and μ must be eliminated from (3) and (4). Denote $\phi^*_i(\lambda_j, \lambda_k, \lambda_l)$ by ϕ^*_i, then elimination of μ from (3) gives

$$(5)\quad \begin{aligned} &(a)\ x_1(\phi^*_3\psi^*_4 - \phi^*_4\psi^*_3) + x_3(\phi^*_4\psi^*_1 - \phi^*_1\psi^*_4) + x_4(\phi^*_1\psi^*_3 - \phi^*_3\psi^*_1) = 0,\\ &(b)\ x_2(\phi^*_3\psi^*_4 - \phi^*_4\psi^*_3) + x_3(\phi^*_4\psi^*_2 - \phi^*_2\psi^*_4) + x_4(\phi^*_2\psi^*_3 - \phi^*_3\psi^*_2) = 0.\end{aligned}$$

These are equations of degree $r+s$ in the λ's. Elimination of the λ's gives a symmetric equation of degree $mn(r+s)$ in x_1, x_3, x_4 and of degree[8] $mn(r+s)$ in x_2, x_3, x_4, hence of degree $mn(r+s)$ in x_1, x_2, x_3, x_4. If however equations (4) and (5) have k points in common which are in the same order $\alpha, \beta, \gamma, \delta$-fold four the four equations, then the resultant is of degree $n(r+s)^2 - k\beta\gamma\delta$ in the coefficients of M; $n(r+s)^2 - k\alpha\gamma\delta$ in the coefficients of N; $mn(r+s) - k\alpha\beta\delta$ in the coefficients of (5_a); $mn(r+s) - k\alpha\beta\gamma$ in the coefficients of (5_b). When $\delta = \gamma$, then the resultant (reduced) is of degree $mn(r+s) - k\alpha\beta\gamma$. Hence

THEOREM 3. *The parametric equations* (3) *in conjunction with* (4) *and the indicated multiplicities represent a ruled symmetric surface R of order $mn(r+s) - k\alpha\beta\gamma$.*

3. *Example of octic ruled surface.* If (3) has the form

$$(6)\quad \begin{aligned} \rho x_1 &= \lambda_2\lambda_3\lambda_4 + \mu\lambda_1,\\ \rho x_2 &= \lambda_1\lambda_3\lambda_4 + \mu\lambda_2,\\ \rho x_3 &= \lambda_1\lambda_2\lambda_4 + \mu\lambda_3,\\ \rho x_4 &= \lambda_1\lambda_2\lambda_3 + \mu\lambda_4,\end{aligned}$$

and

$$(7)\quad \begin{aligned} M &= \Sigma\lambda_i\lambda_k,\\ N &= \Sigma\lambda_i\lambda_k\lambda_l;\end{aligned}$$

$k = 4$ (vertices of coördinate tetrahedrons), $\alpha = 1$, $\beta = 2$, $\delta = \gamma = 2$, then

[8] The resultant of equations (4) and (5) according to the conventional methods is of degree $2mn(r+s)$, but reduces to the degree $mn(r+s)$ after deleting extraneous factors. This is verified by means of the $(1,1)$ correspondence which exists between the curves C_{mnr} and C_{mns} of indicated orders, obtained in a λ-space by the mapping of C_{mn} $(M \times N)$ by means of the ϕ^*'s and ψ^*'s.

the order of R becomes $3\cdot 2\cdot 4 - 4\cdot 2\cdot 2 = 8$. Ordinarily the elimination of the λ's and μ presents great difficulties on account of the labor involved. In this example this task may actually be accomplished as follows: Let $\phi_1, \phi_2, \phi_3, \phi_4$ now stand for the elementary symmetric functions of the variables x, and L_1, L_2, L_3, L_4 for the λ's. Then

$$
(8)\quad
\begin{aligned}
\rho\,\phi_1 &= L_3 + \mu L_1 \\
\rho^2\phi_2 &= L_2L_4 + \mu(L_1L_3 - 4L_4) + \mu^2 L_2 \\
\rho^3\phi_3 &= L_1L_4^2 + \mu L_4(L_1L_2 - L_3L_4) + \mu^2(L_2L_3 - 3L_1L_4) + \mu^3 L_3 \\
\rho^4\phi_4 &= L_4^3 + \mu L_4^2(L_1^2 - 2L_2) \\
&\quad + \mu^2 L_4(L_2^2 - 2L_1L_3 + 2L_4) + \mu^3(L_3^2 - 2L_2L_1) + \mu^4 L_4.
\end{aligned}
$$

When (λ) lies on the intersection C_6 (rational with double points at the A_i) of $L_2 = 0$, $L_3 = 0$, (8) reduces to

$$
(9)\quad
\begin{aligned}
\rho\,\phi_1 &= \mu L_1, \\
\rho^2\phi_2 &= -4\mu L_4, \\
\rho^3\phi_3 &= L_1L_4^2 - 3\mu^2 L_1L_4, \\
\rho^4\phi_4 &= L_4^3 + \mu L_1^2L_4^2 + 2\mu^2L_4^2 + \mu^4 L_4.
\end{aligned}
$$

Eliminating ρ, μ, L_1, L_4 from (9) we obtain the octic ruled surface R:

$$
(10)\quad
\begin{aligned}
&16\phi_1\phi_3\phi_4 - 8\phi_1\phi_2^2\phi_3 - 12\phi_1^2\phi_2\phi_4 + 4\phi_1^2\phi_2^3 - 16\phi_1^2\phi_3^2 \\
&\qquad + 24\phi_1^3\phi_2\phi_3 - 9\phi_1^4\phi_2^2 + 4\phi_2\phi_3^2 = 0.
\end{aligned}
$$

It cuts the unit plane $\phi_1 = 0$ in the conic $(\phi_1 = 0,\ \phi_2 = 0)$ C_2 and contains the three diagonal lines $(\phi_1 = 0,\ \phi_3 = 0)$ as double lines. If $P(\lambda)$ lies on C_6, then by the cubic involution $\rho\lambda'_i = 1/\lambda_i$, $P(\lambda)$ goes into $P'(1/\lambda_1, 1/\lambda_2, 1/\lambda_3, 1/\lambda_4)$ which lies in the intersection of L_1 and L_2, i. e., on C_2. The join of $P'P$ interpreted in S_3 is a generatrix of R. Thus

THEOREM 4. *The locus of joins of corresponding points in the cubic involution T in which C_2 and C_6 are corresponding is an octic symmetric ruled surface R.*

This can be verified by the principle of correspondence: Let g be a generic line in S_3 and (α) the pencil of planes on g. An α cuts C_6 in six points B: to which correspond by T six points B'_i on C_2, which joined to g give six planes α'. Conversely to a plane α' which cuts C_2 in two points B'_i correspond by T two points B_i on C_6, thus determining two planes α. This establishes a $(6, 2)$-correspondence between the planes α and α' with $6 + 2 = 8$ coincidences. Let α^* be such a coincidence, then there lie on this plane two corresponding

points by T, P on C_6 and P' on P_2, such that $P'P$ is a generatrix of R. There are thus 8 such generatrices cutting g. R is of order 8.

4. *Developable ruled surfaces.* Every developable surface which is not a cone has an "edge of regression" or cuspidal curve which may be any space curve. When the developable surface is symmetric, then a tangent of the cuspidal curve is transformed into 24 other such tangents by the G_{24}. From this follows that the cuspidal curve is symmetric and that it may therefore be obtained as the intersection of two symmetric surfaces $f_m = 0$ and $g_n = 0$ of orders m and n respectively. I shall restrict myself to the case in which the intersection of f_m and g_n is a complete irreducible curve C_{mn}. The tangent planes at a point of intersection (y) of the two surfaces, to each f_m and g_n are

$$(11) \qquad \Sigma x_i \frac{\partial f}{\partial y_i} = 0, \qquad\qquad (12) \qquad \Sigma x_i \frac{\partial g}{\partial y_i} = 0.$$

They intersect in a tangent t of the curve of intersection of the two surfaces and as (y) describes the curve of intersections C_{mn}, t describes the developable surface D which is now symmetric. Its order is obtained by eliminating (y) from $f_m = 0$, $g_n = 0$, (11) and (12), which gives for the order of D $mn(n-1) + mn(m-1)$ or

$$(13) \qquad d = mn(m+n-2).$$

The order of the double curve of D, since C_{mn} has supposedly no effective singularities is

$$(14) \qquad \vartheta = \tfrac{1}{2}mn(m+n-2)[mn(m+n-2)-4],$$

according to a well known formula (Cayley). D has a double curve of order $\frac{1}{2}mn(m+n-2)$ (which is always possible, because d is even) in each of the six planes $x_i - x_k = 0$. Hence, outside of these components, the double curve of D contains a residual double curve of order

$$(15) \qquad \sigma = \tfrac{1}{2}mn(m+n-2)[mn(m+n-2)-10].$$

To find the point of intersection of t with the unit plane, we solve (11), (12) and $\Sigma x_i = 0$ for (x). Denoting $\partial f/\partial y_i$ and $\partial g/\partial y_i$ by f_i and g_i respectively

$$(16) \qquad \begin{aligned} \rho x_1 &= f_3g_4 - f_4g_3 + f_4g_2 - f_2g_4 + f_2g_3 - f_3g_2 \\ \rho x_2 &= -(f_3g_4 - f_4g_3 + f_4g_1 - f_1g_4 + f_1g_3 - f_3g_1) \\ \rho x_3 &= f_2g_4 - f_4g_2 + f_4g_1 - f_1g_4 + f_1g_2 - f_2g_1 \\ \rho x_4 &= -(f_2g_3 - f_3g_2 + f_3g_1 - f_1g_3 + f_1g_2 - f_2g_1). \end{aligned}$$

As the point $P(\lambda)$ describes C_{mn}, the point (x) in (16) describes the curve of intersection D' of D with the unit plane. The join of D' with P is a tangent of C_{mn} or a generatrix of D. With μ variable we have

THEOREM 5. *The parametric representation of the developable tangent surface of C_{mn} the curve of intersection of f_m and g_n is given by*

$$(17)\qquad \begin{aligned} \sigma x_1 &= f_3g_4 - f_4g_3 + f_4g_2 - f_2g_4 + f_2g_3 - f_3g_2 + \mu\lambda_1 \\ \sigma x_2 &= -(f_3g_4 - f_4g_3 + f_4g_1 - f_1g_4 + f_1g_3 - f_3g_1) + \mu\lambda_2 \\ \sigma x_3 &= f_2g_4 - f_4g_2 + f_4g_1 - f_1g_4 + f_1g_2 - f_2g_1 + \mu\lambda_3 \\ \sigma x_4 &= -(f_2g_3 - f_3g_2 + f_3g_1 - f_1g_3 + f_1g_2 - f_2g_1) + \mu\lambda_4; \\ & f_m = 0, \quad g_n = 0. \end{aligned}$$

As may be expected these are precisely of the type represented by (3) and (4).

Example. The simplest developable symmetric surface is obtained as the tangent surface of the sextic C_6 of genus four obtained as the intersection of the symmetric cubic and quadric

$$M = \Sigma\lambda_i\lambda_k\lambda_l = 0 \quad\text{and}\quad N = \Sigma\lambda_i\lambda_k, \qquad (i \neq k \neq l = 1, 2, 3, 4).$$

In this case (17) becomes

$$(18)\qquad \begin{aligned} \rho x_1 &= -(\lambda_2 - \lambda_3)(\lambda_3 - \lambda_4)(\lambda_4 - \lambda_2) + \mu\lambda_1 \\ \rho x_2 &= (\lambda_1 - \lambda_3)(\lambda_3 - \lambda_4)(\lambda_4 - \lambda_1) + \mu\lambda_2 \\ \rho x_3 &= -(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_4)(\lambda_4 - \lambda_1) + \mu\lambda_3 \\ \rho x_4 &= (\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_3 - \lambda_1) + \mu\lambda_4 \end{aligned}$$

a symmetric developable surface of order 12, whose equation in the x's is obtained by elimination of ρ, μ, and the λ's from (18) and $M = 0$, $N = 0$. This is omitted on account of the enormous amount of labor which is required.

UNIVERSITY OF ILLINOIS.

ON CIRCLES CONNECTED WITH THREE AND FOUR LINES.*

By J. R. Musselman.

I. Let us consider six points a_i $(i = 1, 2 \cdots 6)$ of a plane (or sphere) which are ordered. A quadratic covariant associated with them has been discussed by the Morleys.[1] Here we study the six points under the condition that the cross-ratio

$$\frac{(a_1 - a_2)(a_3 - a_4)(a_5 - a_6)}{(a_2 - a_3)(a_4 - a_5)(a_6 - a_1)} = \rho, \tag{1.1}$$

where ρ is any real number. Since this cross-ratio is invariant under homographies

$$y = (\alpha x + \beta)/(\gamma x + \delta)$$

we are asking that it be invariant also under antigraphies

$$\bar{y} = (\alpha x + \beta)/(\gamma x + \delta).$$

The significance of (1.1) is easily seen. For the equation

$$\frac{(a_1 - a_2)(a_3 - x)}{(a_2 - a_3)(x - a_1)} = \rho_1$$

represents a circle on the points a_1, a_2 and a_3. Writing similar expressions for the circles $a_3a_4a_5$ and $a_5a_6a_1$, and multiplying the three equations together we see that (1.1) is precisely the condition *that circles $a_1a_2a_3$, $a_3a_4a_5$, $a_5a_6a_1$ have a common point.* Furthermore, the condition that circles $a_2a_3a_4$, $a_4a_5a_6$, $a_6a_1a_2$ have a common point is

$$\frac{(a_2 - a_3)(a_4 - a_5)(a_6 - a_1)}{(a_3 - a_4)(a_5 - a_6)(a_1 - a_2)} = \rho'. \tag{1.2}$$

Now the truth of either one of (1.1) or (1.2) implies the truth of the other, hence we have here a simple proof of the theorem *if the circles $a_1a_2a_3$, $a_3a_4a_5$, $a_5a_6a_1$ meet at a point, say m; then the circles $a_2a_3a_4$, $a_4a_5a_6$, $a_6a_1a_2$ meet at a point, say n.*

If we choose the points a_i so that m is ∞, then a_2, a_4, a_6 lie respectively on the lines a_1a_3, a_3a_5 and a_5a_1. This gives us then the theorem of Miquel—

* Read before the International Congress of Mathematicians, Oslo, July 17, 1936. Received by the Editors, November 30, 1936; revised February 4, 1937.

[1] *Inversive Geometry*, p. 60. See also exercise 6, page 30.

if a point be marked on each side of a triangle, and through each vertex of the triangle and the marked points on the adjacent sides a circle be drawn, the three circles meet at a point. We notice also that the homography which sends a_1 into a_4, a_3 into a_6, and a_5 into a_2 sends the line $a_3a_5\infty$ into the circle a_6a_2n, so that in general a_1a_4, a_3a_6, a_5a_2, and mn are pairs in a homography.

II. In this section we connect the above with the following generalization of an old theorem [2] that *if a point p has images a_2, a_4, a_6 in the sides of a triangle $a_1a_3a_5$, then the circles $a_2a_3a_4$, $a_4a_5a_6$, $a_6a_1a_2$ meet at a point, say n, on the circle $a_1a_3a_5$: and equally the circles $a_1a_2a_3$, $a_3a_4a_5$, $a_5a_6a_1$ meet at a point, say m, on the circle $a_2a_4a_6$.* We give an analytic proof of the first part which enables us to construct the points m and n without the aid of the circles. Let the coördinates of the vertices of the triangle $a_1a_3a_5$ be turns t_i, i.e. $|t_i|=1$. The elementary symmetric functions of the t_i will be denoted by σ_i. The reflection a_2 of the point p in the side a_1a_3 will have as coördinate $t_1+t_3-t_1t_3\bar{p}$. Consider the equation

$$(2.1) \qquad (t_1+t_3)x=\sigma_2-\sigma_3\bar{p}+(t_1t_3-\sigma_3\bar{p})T.$$

For $T=-1$, $t_3t_1^{-1}$, $t_1t_3^{-1}$ respectively, this circle (2.1) passes through the points a_5, a_4, a_6. Also for the turn value

$$T=-\frac{(\sigma_2-\sigma_3\bar{p})(t_5-p)}{(\sigma_1-p)(t_1t_3-\sigma_3\bar{p})}, \qquad x=\frac{\sigma_2-\sigma_3\bar{p}}{\sigma_1-p},$$

a point on the circle $a_1a_3a_5$. Since the coördinate of this point, n, is symmetric in t_1, t_3, t_5, the circles $a_6a_1a_2$ and $a_2a_3a_4$ likewise are on n, which proves the first part of the theorem. The second part follows immediately from the homography which connects the pairs a_1a_4, a_3a_6, a_5a_2, m and n, for it is of period two.

For if p and a_2 are images in the side a_1a_3 then $(a_1-a_2)/(a_2-a_3)$ and $(a_1-p)/(p-a_3)$ must be conjugates. Writing similar expressions for p and a_4, p and a_6 and multiplying the three we obtain

$$\frac{(a_1-a_2)(a_3-a_4)(a_5-a_6)}{(a_2-a_3)(a_4-a_5)(a_6-a_1)}=1.$$

But this is the condition that a_1a_4, a_2a_5, a_3a_6 be pairs in an involution. Naturally m and n belong to this involution. When m is ∞, we have the theorem of Menelaus.

---

[2] This is a generalization of a theorem of Canon, *Nouvelles Annales de Mathématique*, Fourth Series, vol. 8 (1908), p. 480, Problem 2108. See also R. Bouvaist, *ibid.*, vol. 10 (1910), p. 136.

Now if H be the orthocenter of $a_1a_3a_5$, any point on the line Hp will have as coördinate $r=(\sigma_1+\lambda p)/(1+\lambda)$. The coördinate of n will be $(\sigma_2-\sigma_3\bar{r})/(\sigma_1-r)$. Substituting the values for r and $\bar{r}$ in this expression for n we obtain $(\sigma_2-\sigma_3\bar{p})/(\sigma_1-p)$. Since this is independent of the parameter λ, we see that *the point n is a fixed point on $a_1a_3a_5$ for all points on the line Hp.* The Simson line of n is parallel to Hp, so n can be constructed[3] without the use of circles. The coördinate of the point m can be written as

$$m=\sigma_1-p+\frac{p-n}{1-\bar{p}n}, \qquad n=\frac{\sigma_2-\sigma_3\bar{p}}{\sigma_1-p}. \tag{2.2}$$

The line pn cuts the circle $a_1a_3a_5$ at a point L, whose coördinate is

$$L=\frac{p-n}{1-\bar{p}n}. \tag{2.3}$$

Hence if O be the center of the circle $a_1a_3a_5$, the point n can be constructed from the vector relation

$$\overline{Om}=\overline{OH}-\overline{Op}+\overline{OL}. \tag{2.4}$$

III. Casey[4] defines as "twin points" any two points such that the angles subtended by the sides of a triangle at these points are either equal or supplementary. With reference to the triangle $a_1a_3a_5$ the points p and m are twin points. It is also known that any two points at the ends of a diameter of a rectangular hyperbola are twin points with reference to any triangle inscribed in the hyperbola. The theorem in Section II connecting p and m enables us to state two interesting theorems about the rectangular hyperbola. (1) *If we reflect any point p of a rectangular hyperbola in a chord a_1a_3, obtaining the point a_2, then the circle $a_1a_2a_3$ will intersect the hyperbola at m, the diametrically opposite point of p.* (2) *Let A_1, A_2, A_3, P and Q be any five points of a rectangular hyperbola with P and Q at the ends of a diameter, and let P_i, Q_i $(i=1,2,3)$ be the reflections of P and Q respectively in the chords A_2A_3, A_3A_1, A_1A_2; then the circles $A_2A_3P_1$, $A_3A_1P_2$, $A_1A_2P_3$ and $P_1P_2P_3$ will meet at Q and the circles $A_2A_3Q_1$, $A_3A_1Q_2$, $A_1A_2Q_3$ and $Q_1Q_2Q_3$ will meet at P.*

One application to the geometry of the triangle will indicate the importance of this theorem. The isogenic centers or Fermat points F, F' of a triangle $A_1A_2A_3$ lie on the ends of a diameter of the Kiepert hyperbola which also passes through the vertices of the triangle. Hence *if we reflect either Fermat point in the sides A_2A_3, A_3A_1, A_1A_2, obtaining the points B_1, B_2 and B_3 the circles $A_2A_3B_1$, $A_3A_1B_2$, $A_1A_2B_3$, $B_1B_2B_3$ are on the other Fermat point.*

[3] See R. A. Johnson, *Modern Geometry*, page 208, Theorem 329.

[4] *A Sequel to Euclid*, Sixth Edition, page 249.

IV. We consider here the reflections of a point p in four lines. Our coördinate system is chosen so that the parabola which touches the four lines has the form

$$x = \frac{2}{(1-t)^2}.$$

If p be the focus of the parabola, and we reflect it in the triangle formed by the tangents at the points t_1, t_2 and t_3 the coördinate of the point n will be

$$n = \frac{2}{1-\sigma_1} = \frac{2}{1-S_1+t_4},$$

where S_1 is the symmetric function of four points t_i. For each triangle formed by three out of the four lines we shall have a point n, and these four points n lie on the circle C_4, whose equation is

$$x = \frac{2}{1-S_1+t}.$$

The inverse of the focus of the parabola as to this circle is the point

$$z = \frac{2}{1-S_1}.$$

For five lines of a parabola, we can construct five circles as C_4, and the inverse point of the focus in each circle will lie on a circle, namely on

$$x = \frac{2}{1-\sigma_1+t},$$

where now σ_1 refers to five points t_1. The chain runs on indefinitely for lines of a parabola.

When p is a point on the circumcircle of the triangle, we see from (2.3) and (2.4) that L coincides with p and that m is at the orthocenter. Hence *if we reflect the focus of a parabola touching any four lines in the four lines, the four points m, one for each three of the four lines, lie on a line—the directrix of the parabola.*

For p any point in the plane, the coördinate of n is

$$n = \frac{2(p+\bar{p}-2)}{\Pi p + 2(\sigma_1 - 1)}, \qquad (\Pi = 1-\sigma_1+\sigma_2-\sigma_3).$$

The numerator of this fraction equated to zero is the equation of the directrix; the denominator equated to zero gives the orthocenter of the triangle of tangents at the points t_1, t_2, t_3. We thus have the theorem *if we reflect any point*

on the directrix of a parabola in the triangle formed by three of its tangents, the point n coincides with the focus. Naturally n is indeterminate when p is the orthocenter of the triangle of tangents.

V. If we choose as the point p the circumcenter of a triangle $A_1A_2A_3$, then from (2. 2) the coördinate of m can be written as

$$m_4 = \sigma_1 - \sigma_2/\sigma_1$$

where σ_i are symmetric functions of the three symbols t_i. If we take a fourth point A_4 on the circle $A_1A_2A_3$ we can form four triangles from the four points, each determining a point m_i, which four points m_i lie on a circle. For

$$m_4 = \sigma_1 - \sigma_2/\sigma_1 = \frac{S_1^2 - S_2 - t_4S_1}{S_1 - t_4},$$

where S_i are symmetric functions of the four t_i. This, for variable t_4, is the equation of a circle; hence the four points m_i lie on the circle C_5, whose equation may be written as

$$x = \frac{a_2 - a_1t}{a_1 - t}$$

where we write a_2 for $S_1^2 - S_2$ and a_1 for S_1. Now when $t = 0$, $x = a_2/a_1$; when $t = \infty$, $x = a_1$. Therefore, the points a_2/a_1 and a_1 are inverse points as to this circle. Four other pairs of points in the involution set up by this circle are A_4m_4, A_3m_3, A_2m_2, A_1m_1. Hence *the five pairs of points a_2/a_1 and a_1, A_i and m_i are pairs of an involution.* Since

$$\frac{a_2}{a_1} = \frac{a_2 - a_1t_5}{a_1 - t_5}$$

where a_i are the functions S_1 and $S_1^2 - S_2$ written for five symbols t_i, we have shown that if we take five points on a circle, we can determine for each four of the five points a circle such as C_5 and a point such as a_1, *and the inverse point of each a_1 as to its associated circle gives five points which lie on a circle.* This chain can go on indefinitely.

WESTERN RESERVE UNIVERSITY.

ON THE DENSITIES OF INFINITE CONVOLUTIONS.[1]

By Aurel Wintner.

It has been pointed out in a previous paper[2] that, due to certain compactness properties of uniformly bounded functions of uniformly bounded variation, there exists for the "term-by-term" differentiation of infinite convolutions of distribution functions a theorem which does not assume any convergence property of the sequence of derivatives and is, therefore, useful[3] in applications. The present note proves an essential refinement of the theorem in question by showing that, for the "smooth" term of the infinite convolution, the existence of an additional derivative need not be required.[4] In fact, it will be shown that *if at least one term* σ_j *of a convergent infinite convolution* $\sigma_1 * \sigma_2 * \cdots$, *say the term* $\sigma_j = \sigma_1$, *has for* $-\infty < x < +\infty$ *a continuous density of bounded variation, then so does the distribution function* σ *represented by the infinite convolution; and the continuous density of* $\sigma_1 * \cdots * \sigma_n$ *tends, as* $n \to +\infty$, *to that of* $\sigma = \sigma_1 * \sigma_2 * \cdots$ *uniformly in every fixed bounded x-range.*

Let $V(\psi)$ denote the total variation ($\leqq +\infty$) of a function $\psi(x)$, where $-\infty < x < +\infty$. Suppose that the derivatives $\phi'_n(x)$ of a sequence of uniformly bounded, differentiable functions $\phi_n(x)$, $-\infty < x < +\infty$, are such that

(i) $\{\phi'_n(x)\}$ is equicontinuous for $-\infty < x < +\infty$;
(ii) $|\phi'_n(x)| \leqq C$ for some constant C;
(iii) $V(\phi'_n) \leqq c$ for some constant c.

Then[5] the sequence $\{\phi_n(x)\}$ cannot tend to a limit function $\lim \phi_n(x)$ almost everywhere unless $\lim \phi_n(x)$ determines for $-\infty < x < +\infty$ a continuous function which has a uniformly continuous derivative of bounded variation, in which case

$$\phi_n(x) \to \lim_{n\to\infty} \phi_n(x) \quad \text{and} \quad \phi'_n(x) \to (\lim_{n\to\infty} \phi_n(x))'$$

[1] Received January 28, 1937.

[2] A. Wintner, *American Journal of Mathematics*, vol. 57 (1935), pp. 363-366.

[3] Cf. E. R. van Kampen and A. Wintner, *ibid.*, vol. 59 (1937), p. 186 and p. 203.

[4] The assumption made *loc. cit.*[2] was that σ_1 has for $-\infty < x < +\infty$ an absolutely integrable and bounded second derivative. It is clear that this assumption implies the existence of a continuous first derivative which has for $-\infty < x < +\infty$ a bounded variation. Incidentally, a function of bounded variation, which is a derivative, cannot have discontinuities.

[5] This is a corollary of the facts proved *loc. cit.*,[2] pp. 364-365.

hold uniformly in every fixed bounded x-range. Now if $\phi_n = \sigma_1 * \cdots * \sigma_n$, then $\{\phi_n(x)\}$ is uniformly bounded, since ϕ_n is then a distribution function. Furthermore, $\phi_n(x) \to \sigma(x)$ holds almost everywhere, since $\sigma = \sigma_1 * \sigma_2 * \cdots$ is supposed to be a convergent infinite convolution. Hence it is sufficient to prove that the assumption of a derivative $\sigma'_1(x)$, $-\infty < x < +\infty$, of bounded variation implies for the finite convolutions $\phi_n = \sigma_1 * \cdots * \sigma_n$ the existence of densities $\phi'_n = \phi'_n(x)$ which satisfy (i), (ii) and (iii), no matter what the distribution functions $\sigma_2, \sigma_3, \cdots$ may be.

First, if $\omega = \omega(\delta)$ denotes, for a fixed $\delta > 0$, the greatest lower bound of those numbers $\beta > 0$ which have the property that $|\sigma'_1(x^1) - \sigma'_1(x^2)| \leqq \beta$ is a consequence of $|x^1 - x^2| \leqq \delta$ for arbitrary x^1, x^2, then $\omega(\delta) \to 0$ as $\delta \to 0$. In other words, $\sigma'_1(x)$ is not only continuous but uniformly continuous for $-\infty < x < +\infty$. In fact, $\sigma'_1(-\infty) = 0$ and $\sigma'_1(+\infty) = 0$, since the distribution function $\sigma_1(x)$ is supposed to be such that $V(\sigma'_1) < +\infty$. This also implies that $0 \leqq \sigma'_1(x) < V(\sigma'_1)$ for every x.

On placing $\tau_n = \sigma_2 * \cdots * \sigma_n$, it is clear that

$$\phi_n(x) = \phi_n = \sigma_1 * \cdots * \sigma_n = \sigma_1 * \tau_n = \int_{-\infty}^{+\infty} \sigma_1(x-y)\,d\tau_n(y).$$

Since σ'_1 is continuous and bounded, and since τ_n is a distribution function, it follows that ϕ_n has for every x a derivative ϕ'_n which may be obtained by differentiation beneath the integral sign, so that

$$\phi'_n(x) = \int_{-\infty}^{+\infty} \sigma'_1(x-y)\,d\tau_n(y).$$

Now $|x^1 - x^2| \leqq \delta$ implies that $|\sigma'_1(x^1 - y) - \sigma'_1(x^2 - y)| \leqq \omega(\delta)$ for every y, since $|x^1 - x^2| = |(x^1 - y) - (x^2 - y)|$. Hence

$$|\phi'_n(x^1) - \phi'_n(x^2)| \leqq \int_{-\infty}^{+\infty} \omega(\delta)\,d\tau_n(y) = \omega(\delta) \quad \text{whenever} \quad |x^1 - x^2| \leqq \delta.$$

This proves (i), since $\omega(\delta) \to 0$ as $\delta \to 0$. Similarly, since $|\sigma'_1| < V(\sigma'_1)$,

$$|\phi'_n(x)| \leqq \int_{-\infty}^{+\infty} V(\sigma'_1)\,d\tau_n(y) = V(\sigma'_1),$$

so that (ii) is satisfied by $C = V(\sigma'_1)$. In order to prove (iii), notice first

that, since $V(\sigma'_1) < +\infty$, one can write the above representation of $\phi'_n(x)$ in the form

$$\phi'_n(x) = [\tau_n(y)\sigma'_1(x-y)]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \tau_n(y)\,d\sigma'_1(x-y),$$

where, as pointed out above, $\sigma'_1(-\infty) = 0$ and $\sigma'_1(+\infty) = 0$, while $\tau_n(y)$, being a distribution function, is bounded. Hence

$$\phi'_n(x) = -\int_{-\infty}^{+\infty} \tau_n(y)\,d\sigma'_1(x-y), \text{ i.e., } \phi'_n(x) = \int_{-\infty}^{+\infty} \tau_n(x-y)\,d\sigma'_1(y),$$

and so, for arbitrary m and x_i,

$$\sum_{i=1}^{m} |\phi'_n(x_i) - \phi'_n(x_{i-1})| \leqq \int_{-\infty}^{+\infty} \sum_{i=1}^{m} |\tau_n(x_i - y) - \tau_n(x_{i-1} - y)|\,|d\sigma'_1(y)|.$$

Now let $x_0 < \cdots < x_m$. Then $x_0 - y < \cdots < x_m - y$ for every y; hence

$$\sum_{i=1}^{m} |\tau_n(x_i - y) - \tau_n(x_{i-1} - y)| \leqq V(\tau_n) = 1.$$

Consequently,

$$\sum_{i=1}^{m} |\phi'_n(x_i) - \phi'_n(x_{i-1})| \leqq \int_{-\infty}^{+\infty} |d\sigma'_1(y)| = V(\sigma'_1) \text{ whenever } x_0 < \cdots < x_m,$$

and so $V(\phi'_n) \leqq V(\sigma'_1)$; i.e., (iii) is satisfied by $c = V(\sigma'_1)$.

It is clear from the proof that the theorem can be extended to the case of multi-dimensional distributions and also to the case where one requires for $\sigma = \sigma_1 * \sigma_2 * \cdots$ derivatives higher than the first.

THE JOHNS HOPKINS UNIVERSITY.

A REPRESENTATION OF STIELTJES INTEGRALS BY CONDITIONALLY CONVERGENT SERIES.*

By Fritz John.

In an earlier paper the author expressed $\int_0^1 f(x)dx$, where $f(x)$ is a function of bounded variation by a series of the form $\sum_{\nu=1}^{\infty} a_\nu f(\lambda_\nu)$, a_ν and λ_ν being certain constants not depending on f.[1] The a_ν and λ_ν contained an arbitrary rational parameter γ. These expressions for the integral of a function were generalized by H. Rademacher,[2] who assigned to γ arbitrary algebraic values, and also proved the validity of the expansion for all Riemann integrable functions.

In the present paper I am giving a similar representation for Riemann-Stieltjes integrals:

$$\int_1^2 f(x)d\psi(x) = \sum_{\nu=0}^{\infty} a_\nu f(c_\nu) \tag{1}$$

where the a_ν and c_ν are independent of f. The sequences a_ν and c_ν are not uniquely determined by $\psi(x)$. One might expect that the c_ν can be prescribed arbitrarily to a certain extent (e. g. so as to form an everywhere dense set in the interval $1 \leqq x \leqq 2$) and that then coefficients a_ν depending on ψ can be determined such, that (1) holds for all functions f of a certain class. In this paper a special expansion of this sort is given for the case, that f is of bounded variation and that ψ is continuous; the arguments c_ν have the fixed values $\nu/2^{[\log_2 \nu]}$, independent of ψ. The series is in general only conditionally convergent. By introducing $y = \psi(x)$ as variable of integration, one can obtain new expansions $\sum_{\nu=0}^{\infty} a_\nu f(\lambda_\nu)$ for the Riemann integral $\int_1^2 f(y)dy$, the coefficients a_ν and the arguments λ_ν depending on an arbitrary continuous, monotonic function $\psi(x)$.

It would be desirable, to prove the more general theorem, that every continuous linear operator for functions of bounded variation f (the total

* Received December 2, 1936.

[1] "Identitäten zwischen dem Integral einer Funktion und unendlichen Reihen," *Mathematische Annalen*, vol. 110, pp. 718-721.

[2] "Some remarks on F. John's identity," *American Journal of Mathematics*, vol. 58 (1936), pp. 169-176.

variation of f taken as its norm) can be represented by a series of the form $\sum_\nu a_\nu f(c_\nu)$.

In what follows we shall denote with $\bar{x}$ the function defined by

$$\bar{x} = \begin{cases} \dfrac{x}{2^{[\log_2 x]}} & \text{for } x > 0 \\ 1 & \text{for } x = 0. \end{cases}$$

THEOREM. *Let $f(x)$ be of bounded variation and $\psi(x)$ be continuous in $1 \leqq x \leqq 2$. Then*

$$\int_1^2 f(x)\,d\psi(x) = \sum_{\nu=0}^{\infty} a_\nu f(\bar{\nu}),$$

the coefficients a_ν being defined by the following conditions:

(2) $\quad a_{2\nu} = \tfrac{1}{2}a_\nu + \tfrac{1}{2}\big(\psi(\overline{2\nu-2}) - \psi(\overline{2\nu-1})\big)$ for positive integers ν,

(3) $\quad a_\nu = \tfrac{1}{2}\big(\psi(\bar{\nu}) - \psi(\overline{\nu-1})\big)$ for *odd* positive integers ν,

(4) $\quad a_0 = \psi(2) - \psi(1)$.

Proof. We have for $\nu = 0, 1, 2, \cdots$

$$\overline{2\nu} = \bar{\nu}, \qquad 1 \leqq \bar{\nu} < 2. \tag{5}$$

Moreover, as $\psi(x)$ is continuous, there exists for every positive ϵ a $\delta(\epsilon)$ such that $|\psi(x_1) - \psi(x_2)| < \epsilon$ for $1 \leqq x_1 \leqq 2$, $1 \leqq x_2 \leqq 2$ and $|x_1 - x_2| \leqq \delta(\epsilon)$.

Let ν be an odd number $> 2/\delta(\epsilon) + 1$. Then $[\log_2 \nu] = [\log_2 (\nu - 1)]$ and

$$|\bar{\nu} - \overline{\nu-1}| = \frac{1}{2^{[\log_2 \nu]}} \leqq \delta;$$

therefore $|a_\nu| \leqq \epsilon/2$.

Let ν be even and $> 2/\delta(\epsilon) + 1$. Then

$$[\log_2(\nu-2)] = [\log_2(\nu-1)],$$
$$|\overline{\nu-2} - \overline{\nu-1}| = \frac{1}{2^{[\log_2(\nu-1)]}} < \delta,$$
$$|\psi(\overline{\nu-2}) - \psi(\overline{\nu-1})| < \epsilon,$$
$$|a_\nu| = |\tfrac{1}{2}a_{\nu/2} + \tfrac{1}{2}\big(\psi(\overline{\nu-2}) - \psi(\overline{\nu-1})\big)| \leqq \tfrac{1}{2}|a_{\nu/2}| + \epsilon/2.$$

Thus for all $\nu > 2/\delta(\epsilon) + 1$

$$|a_\nu| \leqq \tfrac{1}{2}|a_{[\nu/2]}| + \epsilon/2.$$

Consequently we have, if M_n denotes $\displaystyle\operatorname*{Maximum}_{2^n \leqq \nu < 2^{n-1}} a_\nu$

$$M_{n+1} \leqq \tfrac{1}{2}M_n + \epsilon/_2$$

for sufficiently large n; therefore $\limsup_{n\to\infty} M_n \leqq \epsilon$, and as ϵ was an arbitrary positive number, $\lim_{n\to\infty} M_n = 0$. Hence

$$\lim_{\nu\to\infty} a_\nu = 0. \tag{6}$$

Let for $n = 0, 1, 2, \cdots$ and for $1 \leqq x < 2$

$$s_n(x) = \sum_{2^n \leqq \nu \leqq 2^n x} a_\nu.$$

Then for positive n

$$s_n(x) = \sum_{2^{n-1} \leqq \nu \leqq 2^{n-1}x} (a_{2\nu} + a_{2\nu-1}) - a_{2^n-1} + \theta_n a_{[2^n x]},$$

where

$$\theta_n = \begin{cases} 1 \text{ if } [2^n x] \text{ is odd} \\ 0 \text{ if } [2^n x] \text{ is even.} \end{cases}$$

Now according to (2), (3)

$$a_{2\nu} + a_{2\nu-1} = \tfrac{1}{2}a_\nu + \tfrac{1}{2}(\psi(\overline{2\nu-2}) - \psi(\overline{2\nu-1})) + \tfrac{1}{2}\big(\psi(\overline{2\nu-1}) - \psi(\overline{2\nu-2})\big) = \tfrac{1}{2}a_\nu.$$

Hence

$$s_n(x) = \tfrac{1}{2}s_{n-1}(x) - a_{2^n-1} + \theta_n a_{[2^n x]}. \tag{7}$$

As according to (6) $|a_{2^n-1}|$ and $|a_{[2^n x]}|$ are less than $\epsilon/4$ for sufficiently big n, we have for those n

$$|s_n(x)| \leqq \tfrac{1}{2}|s_{n-1}(x)| + \epsilon/2.$$

From this we may conclude, that

$$\lim_{n\to\infty} s_n(x) = 0 \text{ uniformly in } x \text{ for } 1 \leqq x < 2. \tag{8}$$

It follows moreover from (7), that

$$\sum_{n=1}^{N} s_n(x) = \tfrac{1}{2}\sum_{n=0}^{N-1} s_n(x) - \sum_{n=1}^{N} a_{2^n-1} + \sum_{n=1}^{N} \theta_n a_{[2^n x]}. \tag{9}$$

Now for $n \geqq 1$

$$a_{2^n-1} = \tfrac{1}{2}\big(\psi(\overline{2^n-1}) - \psi(\overline{2^n-2})\big) = \tfrac{1}{2}\big(\psi(\overline{2^n-1}) - \psi(\overline{2^{n-1}-1})\big)$$

$$\sum_{n=1}^{N} a_{2^n-1} = \tfrac{1}{2}\psi(\overline{2^N-1}) - \tfrac{1}{2}\psi(\overline{0});$$

besides for odd $[2^n x]$

$$\theta_n a_{[2^n x]} = \tfrac{1}{2}\big(\psi(\overline{[2^n x]}) - \psi(\overline{[2^n x - 1]})\big)$$
$$= \tfrac{1}{2}\big(\psi(\overline{[2^n x]}) - \psi(\overline{[2^{n-1}x]})\big)$$

and for even $[2^n x]$: $[2^n x] = 2[2^{n-1}x]$

$$\theta_n a_{[2^n x]} = 0 = \tfrac{1}{2}\left(\psi(\overline{[2^n x]}) - \psi(\overline{[2^{n-1}x]})\right).$$

Thus

$$\sum_{n=1}^{N} \theta_n a_{[2^n x]} = \tfrac{1}{2}\left(\psi(\overline{[2^N x]}) - \psi(\overline{[x]})\right) = \tfrac{1}{2}\left(\psi(\overline{[2^N x]}) - \psi(1)\right).$$

Substituting these expressions in (9), we obtain

$$\tfrac{1}{2}\sum_{n=0}^{N} s_n(x) = -\tfrac{1}{2}s_N(x) + s_0(x) - \tfrac{1}{2}\psi(\overline{2^N-1}) + \tfrac{1}{2}\psi(\bar{0}) - \tfrac{1}{2}\psi(1) + \tfrac{1}{2}\psi(\overline{[2^N x]}).$$

Now

$$\begin{aligned} s_0(x) &= a_1 = \tfrac{1}{2}\left(\psi(1) - \psi(\bar{0})\right), \\ \psi(\overline{2^N-1}) &= \psi\left(2 - \frac{1}{2^{N-1}}\right), \\ \lim_{N\to\infty} \psi(\overline{2^N-1}) &= \psi(2), \\ \psi(\overline{[2^N x]}) &= \psi\left(\frac{[2^N x]}{2^N}\right) \end{aligned}$$

$\lim_{N\to\infty} \psi(\overline{[2^N x]}) = \psi(x)$ uniformly in x for $1 \leqq x < 2$. Hence using (8) it follows that

$$\sum_{n=0}^{\infty} s_n(x) = \psi(x) - \psi(2) \text{ uniformly in } x \text{ for } 1 \leqq x < 2. \tag{10}$$

Let the step function $e(y, x)$ be defined by

$$e(y, x) = \begin{cases} 1 & \text{if } y \leqq x \\ 0 & \text{if } y > x. \end{cases}$$

Then for $1 \leqq x < 2$

$$\begin{aligned} \sum_{\nu=1}^{n} a_\nu e(\bar{\nu}, x) &= \sum_{\substack{\nu=1 \\ 1 \leqq \bar{\nu} \leqq x}}^{n} a_\nu \\ &= \sum_{\mu=0}^{[\log_2 n]-1} \sum_{2^\mu \leqq \nu \leqq 2^\mu x} a_\nu + \sum_{2^{[\log_2 n]} \leqq \nu \leqq \min(n,\, 2^{[\log_2 n]} x)} a_\nu \\ &= \sum_{\mu=0}^{[\log_2 n]-1} s_\mu(x) + \begin{cases} s_{[\log_2 n]}(x) & \text{if } \bar{n} > x \\ s_{[\log_2 n]}(n) & \text{if } \bar{n} \leqq x. \end{cases} \end{aligned}$$

Because of (8) and (10) it follows that

$$\sum_{\nu=1}^{\infty} a_\nu e(\bar{\nu}, x) = \psi(x) - \psi(2) \text{ uniformly in } x \tag{11}$$

for $1 \leqq x < 2$. As $\lim_{x\to 2} e(\bar{\nu}, x) = 1$ and $\lim_{x\to 2} \psi(x) = \psi(2)$, (11) holds for

$1 \leqq x \leqq 2$; i. e.

$$\sum_{\nu=1}^{\infty} a_\nu = 0. \tag{12}$$

Let now $f(x)$ be of bounded variation in $1 \leqq x \leqq 2$.

We first assume, that f is also continuous. Then

$$f(y) = -\int_1^2 e(y, x)\, df(x) + f(2).$$

Consequently

$$\sum_{\nu=1}^{n} a_\nu f(\bar{\nu}) = -\int_1^2 \sum_{\nu=1}^{n} a_\nu e(\bar{\nu}, x)\, df(x) + f(2) \sum_{\nu=1}^{n} a_\nu.$$

As $\sum_{\nu=1}^{\infty} a_\nu e(\nu, x)$ converges uniformly in x and $f(x)$ is of bounded variation, it follows using (11) and (12)

$$\begin{aligned}
\sum_{\nu=1}^{\infty} a_\nu f(\bar{\nu}) &= -\int_1^2 \big(\psi(x) - \psi(2)\big)\, df(x) \\
&= -\int_1^2 \psi(x)\, df(x) + \psi(2)\,\big(f(2) - f(1)\big) \\
&= \int_1^2 f(x)\, d\psi(x) - f(2)\psi(2) + f(1)\psi(1) \\
&\qquad + \psi(2)f(2) - \psi(2)f(1) \\
&= \int_1^2 f(x)\, d\psi(x) + f(1)\,\big(\psi(1) - \psi(2)\big).
\end{aligned}$$

This proves our theorem (cf. the definition of a_0) for the case that f is continuous.

Let $f(x)$ be of bounded variation in $1 \leqq x \leqq 2$ and have discontinuities at the points ξ_μ $(\mu = 1, 2, \cdots)$. Let

$$f(\xi_\mu + 0) - f(\xi_\mu - 0) = -c_\mu, \qquad f(\xi_\mu) - f(\xi_\mu - 0) = d_\mu.$$

Let moreover $\chi(x, y)$ be defined by

$$\chi(x, y) = \begin{cases} 0 \text{ if } x \neq y \\ 1 \text{ if } x = y. \end{cases}$$

Then $f(x) = f_1(x) + f_2(x) + f_3(x)$, where $f_1(x)$ is continuous and of bounded variation,

$$f_2(x) = \sum_{\mu=1}^{\infty} c_\mu e(x, \xi_\mu), \qquad f_3(x) = \sum_{\mu=1}^{\infty} d_\mu \chi(\xi_\mu, x),$$

the series $\sum_{\mu=1}^{\infty} |\, c_\mu \,|$ and $\sum_{\mu=1}^{\infty} |\, d_\mu \,|$ being convergent. It is sufficient to prove our theorem for $f = f_1$, $f = f_2$, $f = f_3$ separately. For $f = f_1$ it follows from our

previous considerations. As (1) holds according to (11) for $f(x) = e(x, \xi_\mu)$ and $\sum_{\nu=1}^{\infty} a_\nu e(\bar{\nu}, \xi_\mu)$ converges uniformly in μ and $\sum_\mu c_\mu$ converges absolutely, it follows, that (1) holds as well for $f = f_2(x)$. In order to prove (1) for $f = f_3(x)$, it is only necessary to prove, that

$$\sum_{\nu=0}^{\infty} a_\nu \chi(\nu, \xi_\mu) = 0 \tag{13}$$

uniformly in μ. But, as

$$\chi(\bar{\nu}, x) = \lim_{h \to +0} \big(e(\nu, x) - e(\nu, x - h) \big)$$

it follows from (11), that

$$\sum_{\nu=1}^{\infty} a_\nu \chi(\bar{\nu}, x) = \psi(x) - \psi(x - 0) = 0$$

uniformly in x for $1 \leqq x \leqq 2$.

Thus (1) is proved generally for any $f(x)$ of bounded variation.

Remarks. If there is an ϵ such that $|f(x)| > \epsilon$ for all x in $1 \leqq x \leqq 2$, then the series $\sum_{\nu=0}^{\infty} a_\nu f(\bar{\nu})$ is certainly not absolutely convergent, unless $\psi(x)$ is a constant.

For if that series would be absolutely convergent, then $\sum_\nu |a_\nu|$ would be convergent; consequently according to (11), (13)

$$\begin{aligned}
\psi(x) - \psi(2) &= \sum_{\nu=1}^{\infty} a_\nu e(\bar{\nu}, x) = \sum_{\nu \text{ odd}}^{\nu} \sum_{s=0}^{\infty} a_{2^s\nu} e(\overline{2^s\nu}, x) \\
&= \sum_{\nu \text{ odd}}^{\nu} \sum_{s=0}^{\infty} a_{2^s\nu} e(\bar{\nu}, x) \\
&= \sum_{\nu \text{ odd}}^{\nu} e(\bar{\nu}, x) \sum_{s=0}^{\infty} a_{2^s\nu} \\
&= \sum_{\nu \text{ odd}}^{\nu} e(\bar{\nu}, x) \sum_{\mu=0}^{\infty} a_\mu \chi(\bar{\mu}, \bar{\nu}) = 0.
\end{aligned}$$

II. Let $\psi(x)$ be monotonically increasing and continuous and let $\psi(2) = 2$, $\psi(1) = 1$. Then we obtain by substituting $t = \psi(x)$ in (1) the identity

$$\int_1^2 f(t)\,dt = \sum_{\nu=0}^{\infty} a_\nu f(\lambda_\nu)$$

valid for every function f of bounded variation, a_ν being defined by (2), (3), (4) and λ_ν denoting $\psi(\bar{\nu})$.

UNIVERSITY OF KENTUCKY.

NOTE ON THE DEFINITION OF FIELDS BY INDEPENDENT POSTULATES IN TERMS OF THE INVERSE OPERATIONS.*

By David G. Rabinow.

1. Introduction. The concept of a field involves two operations which are usually called addition and multiplication. When the definition of the field is given in terms of these operations, we say it is defined for the direct operations.[1] The inverse operations of these direct operations may be called subtraction and division. It is possible to define a field in terms of these inverse operations.[2] The inverse operation of multiplication, division, when used as the fundamental operation, allows us to define multiplication in two essentially distinct ways. The first involves the fact that $1/1/a = a$ under certain restrictions on a. We can then define $ab = a(1/1/b)$. This is the method used in the paper referred to in footnote 2. The second method is probably the more fundamental in that it is analogous to the definition of division in terms of multiplication. We shall develop the definition of a field in terms of addition and division where our treatment of division shall be this second method. Instead of using subtraction we shall use addition, since this will simplify the proofs somewhat and since the use of subtraction has been completely discussed in the paper referred to in footnote 2.

2. Postulates for a field and theorems deducible from them. Let us consider the following set of postulates in connection with the base $(K, +, \mathrm{o})$ where K is a class of elements $a, b, c, \cdots$ and $+, \mathrm{o}$ are binary operations.

Postulate 1. a in K and b in K imply $a + b$ in K.

Postulate 2. If $a, b, c, a + b, b + c, (a + b) + c, a + (b + c)$ are in K, then $(a + b) + c = a + (b + c)$.

Postulate 3. There exists in K at least one element Z such that $a + Z = a$ for all a in K.

* Received December 28, 1936.

[1] E. V. Huntington, "Note on the definition of abstract groups and fields by sets of independent postulates," *Transactions of the American Mathematical Society*, vol. 6 (1905), pp. 181-193.

[2] D. G. Rabinow, "Independent sets of postulates for abelian groups and fields in terms of the inverse operations," *American Journal of Mathematics*, vol. 59 (1937), pp. 211-224.

Postulate 4. For each element a in K there exists at least one element a' in K such that $a + a' = Z$.

Postulate 5. a in K and b in K and $b \neq Z$ imply aob in K.

Postulate 6. If a, b, c, aob, aoc, $(aob)oc$, $(aoc)ob$ are in K, then $(aob)oc = (aoc)ob$.

Postulate 7. If $a, b, c, a + b, aoc, boc, (a + b)oc$ and $aoc + boc$ are in K, then $(a + b)oc = aoc + boc$.

Postulate 8. a in K and b in K and $a \neq Z$ imply the existence of an unique element x in K such that $xoa = b$.

Postulate 9. If a, b, aoa, bob are in K, then $aoa = bob$.

Postulate 10. There exist at least two distinct elements in K.

Note 1. If we wish we may remove the uniqueness requirement in Postulate 8 and insert as an additional postulate either Lemma 2 or Lemma 3 proven below. However for compactness and for simplification of independence proofs, as well as for reasons which will become apparent in Section 4, the present form of Postulate 8 is desirable.

Note 2. Throughout the subsequent work the terms Postulate, Theorem, Lemma and Definition will be referred to respectively by P, T, L, and D. Those theorems deducible from Postulates 1 through 4 shall be assumed as known and will be referred to as G.

Lemma 1. *If $b \neq Z$, then $Zob = Z$.*

Let a be any element in K. Hence since $b \neq Z$, we have by P5, P3, and P7, $aob = (a + Z)ob = aob + Zob$. Whence by G, $Zob = Z$.

Lemma 2. *If $a \neq Z$, $b \neq Z$, then $aob \neq Z$.*

Since $b \neq Z$, there exists by P8 an unique element x such that $xob = Z$. Hence by L1, x must be Z. If $aob = Z$, then a must be Z. But this is a contradiction. Hence $aob \neq Z$.

Lemma 3. *If a, b, c, aoc, boc are in K and if $c \neq Z$ and if $aoc = boc$, then $a = b$.*

By P4 there exists an element a' such that $a + a' = Z$. By P5 $a'oc$ is in K. Hence by P1 $a'oc + boc = a'oc + aoc$. Whence by P7 $(a' + b)oc = (a + a')oc = Zoc$ by P4. Hence by L2, $a' + b = Z$. Therefore by G $a = b$.

At this stage we are in the position to define:

Definition 1. There exists an unique element $U = aoa \neq Z$. For, by P10 there exists in K at least one element $a \neq Z$. Hence by P5 aoa is in K and by L2 $aoa \neq Z$. By P9 this defines the unique element $U = aoa = bob$.

LEMMA 4. *If under the conditions of* P8 *the element* b *is also not equal to* Z, *then* $xob = a$.

For, $b = xoa$. Hence by P5 $(xoa)ob = bob = U$ by D1. Whence by P6 $(xob)oa = U = aoa$ since $a \neq Z$. Therefore by L3 $xob = a$.

We can now define the product (written ab) of any two elements a and b as follows:

Definition 2.1. $b = Z$, then $ab = Z$.

Definition 2.2. $b \neq Z$, then ab shall be the element x of P8, that is the element x satisfying the equation $xob = a$. From these definitions we have immediately

THEOREM 1. a *in* K *and* b *in* K *imply* ab *in* K.

THEOREM 2. *If* a, b, ab, ba *are in* K, *then* $ab = ba$.

Case I. $b = Z$, then by D2.1 $ab = Z$. If $a = Z$, then by D2.1 $ba = Z$. If $a \neq Z$, then by D2.2 and L2 $ba = Z$.

Case II. $a \neq Z$, $b \neq Z$. By D2.2 $ab = x$ where $xob = a$. By D2.2 $ba = y$ where $yoa = b$ and by L4 $yob = a$. Hence by L3 $x = y$.

THEOREM 3. *If* $a, b, c, ab, bc, (ab)c, a(bc)$ *are in* K, *then* $(ab)c = a(bc)$.

Case I. If $a = Z$ or $b = Z$ or $c = Z$, the theorem follows as in Case I of T2.

Case II. $a \neq Z$, $b \neq Z$, $c \neq Z$. By D2.2 let $(ab)c = w$ where $woc = ab$. Hence by D2.2 $(woc)ob = a$. Similarly let $p = a(bc) = (bc)a$ by T2 where $poa = bc$ and hence $(poa)oc = b$ by D2.2. By P6 $b = (poa)oc = (poc)oa$. Hence by L4 $(poc)ob = a$. Therefore $(woc)ob = (poc)ob$ and by repeated application of L3 $p = w$.

THEOREM 4. *If* $a, b, c, a + b, ac, bc, (a + b)c$ *and* $ac + bc$ *are in* K, *then* $(a + b)c = ac + bc$.

Case I. $c = Z$. By D2.1 $(a + b)c = ac = bc = Z$; whence by P3 $(a + b)c = ac + bc$.

Case II. $c \neq Z$. By D2.2 let $(a + b)c = p$ where $poc = a + b$. Also let $ac = x$ where $xoc = a$ and let $bc = y$ where $yoc = b$. Then by P1 $a + b = xoc + yoc = (x + y)oc$ by P7. Hence by L3 $p = x + y$.

THEOREM 5. *If* $a, b, c, a + b, ca, cb, c(a + b)$ *and* $ca + cb$ *are in* K, *then* $c(a + b) = ca + cb$.

This theorem follows immediately from T4 and T2.

THEOREM 6. *There exists an unique element* $u \neq Z$ *such that* $au = ua = a$ *for all* a *in* K.

Consider the element $U \neq Z$ defined in D1.

Case I. $a = Z$. By D2. 1 $Ua = Z$ and by T2 $Ua = aU$.

Case II. $a \neq Z$. Let $Ua = p$ where $poa = U = aoa$ by D2. 2 and D1. Hence by L3 $p = a$. By T2 $aU = Ua$. Now suppose there exists another U' such that $aU' = U'a = a$. From $U'a = a$ we have by D2. 2 $aoa = U'$. But $aoa = U$. Hence $U = U'$, and U is unique.

THEOREM 7. *For any elements* a *and* b *in* K *where* $a \neq Z$, *there exists an unique element* x *such that* $xa = ax = b$.

Let $x = boa$. Then by D2. 2 $p = xa = (boa)a$ where $poa = boa$. Hence by L3 $p = b$ and by T2 $ax = xa$. Now suppose there exists another element x' such that $ax' = x'a = b$. Then by D2. 2 $x' = boa$. Hence $x' = x$.

THEOREM 8. *If* a, b, $a + b$, $b + a$ *are in* K, *then* $a + b = b + a$. Let d be any element of $K \neq Z$ and let D be the element x of T7 such that $dD = U$. By T4 and T5

$$(a + b)(d + d) = a(d + d) + b(d + d) = ad + ad + bd + bd.$$

Likewise by T4 and T5

$$(a + b)(d + d) = (a + b)d + (a + b)d = ad + bd + ad + bd.$$

Therefore

$$ad + ad + bd + bd = ad + bd + ad + bd.$$

Hence by P1, P2, P3, and P4 $ad + bd = bd + ad$ or by T5 $(a + b)d = (b + a)d$. Multiplying by D and using T1, T3, T7 and T6 we have $a + b = b + a$.

But P1, P2, P3, P4, T1, T2, T3, T4, T5, T6, T7, T8 are the postulates for a field in terms of the direct operations of addition and multiplication (Huntington). Hence any system $(K, +, \mathrm{o})$, which satisfies Postulates 1 through 10, is a field with respect to the direct operations of addition and multiplication. Furthermore from Theorem 7 we see immediately that the operation o is the inverse operation of multiplication. To complete the proof that our set of postulates is both a necessary and sufficient set to define a field we must show that from P1, P2, P3, P4, T1, T2, T3, T4, T5, T6, T7, T8 we can deduce P5, P6, P7, P8, P9, P10. For this purpose we define the operation o as follows:

Definition 3. If a is in K and b is in K and if $a \neq Z$, then $x = boa$ shall be the element x in T7 satisfying $xa = b$.

LEMMA 5. *If a, b, c, ac, bc are in K and if $c \neq Z$ and if $ac = bc$, then $a = b$.*

For, by T7 the element c' exists such that $cc' = U$. Then by T1 $(ac)c' = (bc)c'$ or by T4, T7, T6 $a = b$.

THEOREM 9. *a in K, b in K and $b \neq Z$ imply aob in K.* (P5)
Follows immediately from D3.

THEOREM 10. *There exist at least two distinct elements in K.* (P10)
These are the elements Z and u whose existence is postulated in P3 and T6.

THEOREM 11. *a, b, aoa, bob in K imply $aoa = bob$.* (P9)
By T6 $ua = a$ for all a. If $a \neq Z$, then by D3 $u = aoa$ for all a for which aoa is in K.

THEOREM 12. *a in K, b in K, and $a \neq Z$ imply the existence of an unique element x in K such that $xoa = b$.* (P8)

Take $x = ba$. Then by D3 $xoa = b$ if $a \neq Z$. Now suppose there exists another element x' such that $x'oa = b$. Then by D3, $x' = ba$. Hence $x' = x$.

THEOREM 13. *If a, b, c, $a + b$, $(a + b)oc$, aoc, boc, $aoc + boc$ are in K, then $(a + b)oc = aoc + boc$.* (P7)

By P1 and T9 $(a + b)oc$, aoc, boc are in K if $c \neq Z$. By T7 there exists an element x such that $xc = a + b$. Also by T7 there exists elements y and w such that $yc = a$ and $wc = b$. Hence $yc + wc = a + b = xc = (y + w)c$ by P1 and T4. Therefore by L5 $x = y + w$ and the theorem follows by D3.

THEOREM 14. *If a, b, c, aob, aoc, $(aob)oc$, $(aoc)ob$ are in K, then $(aob)oc = (aoc)ob$.* (P6)

Case I. $a = Z$. By D3 $(aob)oc = (aoc)ob = Z$ by T7.

Case II. $a \neq Z$. By T9 aob, aoc, $(aob)oc$, $(aoc)ob$ are in K if $b \neq Z$ and $c \neq Z$. Let $(aob)oc = x$ and $(aoc)ob = w$. By D3 $xc = aob$. Take $aob = y$ where by D3 $yb = a$. Likewise $wb = aoc$ and $aoc = p$ where $pc = a$. Since $xc = y$, then $xcb = yb = a$ and since $wb = p$, then $wbc = pc = a$. Hence $(xc)b = (wb)c = (wc)b$ by T1, T2, T3. Hence by repeated applications of L5 $x = w$.

From the above we conclude that the set of Postulates 1 through 10 is both necessary and sufficient to define a field.

3. Independence of the postulates. The postulates are examined for independence by exhibiting examples of systems $(K, +, \mathrm{o})$ which fail to satisfy the correspondingly numbered postulates but satisfy the remaining postulates.

Example 1) K is the class of two elements $0, 1$ with $a + b$ and $a \mathrm{o} b$ satisfying the following multiplication tables. The elements u, r, and s are elements not in K.

$a+b$	0	1
0	0	u
1	1	0

$a \mathrm{o} b$	0	1
0	r	0
1	s	1.

Example 2) K is the class of all rational numbers, positive, negative, and zero. $a + b = a + 2b$. $a \mathrm{o} b = a/b$.

Example 3) K is the class of all positive rational numbers. $a + b = a + b$. $a \mathrm{o} b = a/b$.

Example 4) K is the class of all positive rational numbers including zero. $a + b = a + b$. $a \mathrm{o} b = a/b$.

Example 5) K is the class of all integers, positive, negative, and zero. $a + b = a + b$. $a \mathrm{o} b = a/b$.

Example 6) K is the class of hypercomplex numbers of the form $\pi 1 + \omega i + \rho j$ where π, ω, ρ are rational numbers, positive, negative, and zero. $a + b = a + b$.

$$a \mathrm{o} b = \frac{1}{\pi_2{}^2 + \rho_2{}^2 + (\omega_2 - \rho_2)^2} \times (\pi_1 1 + \omega_1 i + \rho_1 j)(\pi_2 1 + \omega_2 i + \rho_2 j),$$

where the product of the coefficients shall be the ordinary product of rational numbers and the "units" shall follow the table.

	1	i	j
1	1	$-i$	$-j$
i	i	1	-1
j	j	-1	2.

Example 7) K is the class of all integers, positive, negative and zero. $a + b = a + b$. $a \mathrm{o} b = a - b$.

Example 8) K is the class of all integers, positive, negative and zero. $a + b = a + b$. $a \mathrm{o} b = 0$.

Example 9) K is the class of all rational numbers, positive, negative and zero. $a + b = a + b$. $aob = ab$.

Example 10) K is the class consisting of the element 0 only. $a + b = a + b$. aob is undefined.

4. The concept of operational invariance. Let us consider the Postulates 1 through 8 inclusive. It is to be observed that the commutative postulate, that is $aob = boa$, cannot be deduced from these postulates. An independence example for this postulate would be: K is the class of all rational numbers, positive, negative and zero. $a + b = a + b$ and $aob = a$. We note further that if the commutative postulate is added to the set of Postulates 1 through 8, we obtain the definition of a field in terms of the direct operations, provided that we define $aoZ = Z$. In other words we have found a set of postulates (P1 through P8) which defines a field in terms of either the direct or the inverse operations depending on what additional postulates we desire to add to it. This naturally suggests the following problem: Suppose we consider the set P1 through P8 as a distinct set of postulates involving the operations $+$ and o. Let us further consider the inverse operation of o, which may be defined by means of P8. Call this operation X. Replace o, wherever it occurs in the set P1 through P8, by X. We now have a new set P1′ through P8′. The problem is can we from P1 through P8 deduce P1′ through P8′? The purpose of this section of the paper is to prove that this is true. Any system $(K, +, \mathrm{o})$, which has this property of replacing o by X, is to be defined as an operationally invariant system with respect to the operation o. It is clear that a field does not have this property but that there exists a subset of the postulates of the field, namely P1 through P8, which does. (This concept of operational invariance may obviously be extended to any type of system in which an inverse may be defined.) By P8 there exists a unique element x such that $xob = a$ if $b \neq Z$. This enables us to make the following definition:

Definition 4. $x = aXb$ if x is the element satisfying P8 when $b \neq Z$, that is, if x is the element such that $xob = a$. This proves

THEOREM 15. *a in K, b in K and $b \neq Z$ imply aXb in K.*

THEOREM 16. *If a, b, c, $a + b$, $(a + b)Xc$, aXc, bXc and $aXc + bXc$ are in K, then $(a + b)Xc = aXc + bXc$.*

By T15 aXc, bXc and $(a + b)Xc$ are in K if $c \neq Z$. Let $(a + b)Xc = w$ where by D4 $woc = a + b$. Also let $aXc = p$ and $bXc = q$ where again by D4 $poc = a$ and $qoc = b$. Hence by P1 $poc + qoc = a + b$ or by P7

$(p+q)oc = woc$. Hence $p+q=w$ by L3. (It is to be noted that L1, L2 and L3 are still true since their proofs depended only on P1 through P8.)

THEOREM 17. *a in K and b in K and $a \neq Z$ imply the existence of an unique element w such that $wXa = b$.*

Take $w = boa$. Since $a \neq Z$, then by T15 $(boa)Xa = r$ where r is in K. Hence by D4 $roa = boa$. Whence by L3 $r = b$. Furthermore the element w must be unique since by D4 $w = boa$ which by P5 is uniquely determined by a and b.

LEMMA 6. *If a, b, aXb and $(aXb)ob$ are in K, then $(aXb)ob = a$.*

Let $w = aXb$. By D4 this means $wob = a$.

THEOREM 18. *If a, b, c, aXb, aXc, $(aXb)Xc$ and $(aXc)Xb$ are in K, then $(aXb)Xc = (aXc)Xb$.*

To satisfy the hypothesis of the theorem we see from D4 that $b \neq Z$ and $c \neq Z$. Take $(aXb)Xc = w$ where $woc = aXb$ by D4. Take also $(aXc)Xb = p$ where $pob = aXc$ by D4. Hence by P5 $(woc)ob = (aXb)ob$ and $(pob)oc = (aXc)oc$. By L6 $(aXb)ob = a$ and $(aXc)oc = a$. Hence $(woc)ob = (poc)ob$. Whence by L3 $p = w$.

5. Consistency of the postulates. To show that the set of postulates for a field is consistent we exhibit the set of all rational numbers with ordinary addition and division as our operations. To show the consistency of the set of postulates considered in section 4 we may take the set of all rational numbers with ordinary addition and either ordinary multiplication or ordinary division.

HARVARD UNIVERSITY.

A CORRECTION.

In my paper referred to in footnote 2, one of the postulates is incorrectly stated and several of the independence examples need to be restated.

On page 215, Postulate 18. 2 should be: If a^* exists, then $a^* \neq z$ (provided $a \neq z$, $U \neq z$).

On page 223, Example 12 should be: K is the class of all positive rational numbers excluding zero. $a - b = a + b$. $aob = a/b$.

On page 223, Example 13 should have $a - b = | a - b |$.

On page 223, Example 17 should be: K is the class of all rational numbers, positive, negative, and zero. $a - b = a - b$. $aob = ab$.

On page 223, Example 18. 2 should have $aob = a/b$ except $1/a = 0$, but $1/1 = 1$.

THE REPRESENTATION OF INTEGERS AS SUMS OF VALUES OF CUBIC POLYNOMIALS. II.*

By R. D. James.

1. Introduction. In a previous paper under the same title [1] the author proved the following result.

THEOREM 1. *Let s be an integer* $\geqq 9$ *and let* $P(x)$ *be a polynomial of the form*

$$P(x) = a(x^3 - x)/6 + b(x^2 - x)/2 + cx, \tag{1.1}$$

where a, b, c, *are integers without a common factor, and* $a \not\equiv 4c \pmod 8$. *Then every sufficiently large integer is a sum of nine values of* $P(x)$.

The condition $a \not\equiv 4c \pmod 8$ was an artificial one which could not be removed at the time. In the present paper we shall show that Theorem 1 is true without the restriction $a \not\equiv 4c \pmod 8$. The method of proof was suggested by two recent papers by L. K. Hua.[2] The new idea introduced by him may be explained briefly in the following way. If $\Phi(x)$ is a polynomial of degree k with integral coefficients, an integer θ was defined in § 2, I to be the highest power of a prime p which divided every coefficient of $\Phi'(x)$. Hua defines θ to be the highest power of a prime p for which $p^\theta | \Phi'(x)$ for all integers x. It may be shown that the two definitions are equivalent in the case of cubic polynomials except when $p = 2$. It is this difference when $p = 2$ which enables us to avoid the restriction $a \not\equiv 4c \pmod 8$ in Theorem 1.

The results obtained by Hua in the second of the papers to which reference was made above are correct, but there is an error at the beginning of § 16, page 45. The proofs which he gives are therefore not complete. He makes the following statement: — "Let θ be the highest power of a prime p such that $\Phi'(h) \equiv 0 \pmod{p^\theta}$ for all integers x." The $\Phi(h)$ which he is using is an integral-valued polynomial and not necessarily one with integral coefficients. Hence $\Phi'(h)$ need not be an integer and the congruence $\Phi'(h) \equiv 0$

* Received February 18, 1937.

[1] *American Journal of Mathematics*, vol. 56 (1934), pp. 303-315. This paper will be referred to as I.

[2] *American Journal of Mathematics*, vol. 58 (1936), pp. 553-562; *Journal of the Chinese Mathematical Society*, vol. 1 (1936), pp. 23-61.

(mod p^θ) is meaningless. In Lemmas 1 and 2 we shall show how to avoid this difficulty.

2. The proof of Hua's results. We first introduce the notation to be used. Let

$$P(x) = \sum_{j=0}^{k} a_j \frac{x(x+1)\cdots(x+j-1)}{j!},$$

where the a_j are integers. Let d be the least common multiple of the denominators of

$$a_0,\ a_1,\ a_2/2!,\ a_3/3!,\cdots,a_k/k!.$$

If the canonical product of an integer n is $n = p_1^{l_1}\cdots p_r^{l_r}$, and if $p_i^{a_i} \mid d$, $p_i^{a_i+1} \nmid d$ for $i = 1, 2, \cdots, r$, we define n^* by the equation $n^* = p_1^{l_1+a_1}\cdots p_r^{l_r+a_r}$.

Let $\Phi(x) = d\,P(x)$ so that $\Phi(x)$ is a polynomial with integral coefficients. For every prime p let θ be the highest power of p for which $\Phi'(x) \equiv 0 \pmod{p^\theta}$ for every integer x. Let $P_0(x) = p^{-\theta}\Phi'(x)$. Let $M(m)$ denote the number of solutions of

$$\sum_{\nu=1}^{s} P(x_\nu) \equiv n \pmod{m}, \qquad 0 \leqq x_\nu < m^*. \tag{2.1}$$

For $m = p^l$ let $N(p^l)$ denote the number of solutions of (2.1) in which not every $P_0(x_\nu)$ is divisible by p.

LEMMA 1. (Hua, Lemma 35). *If* $d = p^t D$, *where* $(p, D) = 1$, $l \geqq \max(2\theta + 2 - t, \theta + 2)$ *and* $x = y + dzp^{l-\theta-1}$, *then*

$$P(x) \equiv P(y) + zp^{l-1}P_0(y) \pmod{p^l},$$
$$P_0(x) \equiv P_0(y) \pmod{p}.$$

Proof. If we expand $\Phi(x) = \Phi(y + dzp^{l-\theta-1})$ by Taylor's Theorem we obtain

$$\begin{aligned}\Phi(x) &= \Phi(y) + zdp^{l-\theta-1}\Phi'(y) + \sum_{j=2}^{k}(zdp^{l-\theta-1})^j \frac{\Phi^{(j)}(y)}{j!}\\ &= \Phi(y) + zDp^{t+l-\theta-1}\Phi'(y) + \sum_{j=2}^{k}(zDp^{t+l-\theta-1})^j \frac{\Phi^{(j)}(y)}{j!}.\end{aligned}$$

Since $(t + l - \theta - 1)j \geqq 2(t + l - \theta - 1) \geqq t + l$ we have

$$\begin{aligned}\Phi(x) &\equiv \Phi(y) + zDp^{t+l-\theta-1}\Phi'(y) \pmod{p^{t+l}},\\ dP(x) &\equiv dP(y) + zDp^{t+l-1}P_0(y) \pmod{p^{t+l}},\\ p^tP(x) &\equiv p^tP(y) + zp^{t+l-1}P_0(y) \pmod{p^{t+l}},\\ P(x) &\equiv P(y) + zp^{l-1}P_0(y) \pmod{p^l}.\end{aligned}$$

This proves the first result of the lemma and the second follows in a similar way.

LEMMA 2. (Hua, Lemma 36). *If* $l \geqq \max(2\theta + 2 - t, \theta + 2)$ *then* $N(p^l) = p^{s-1}N(p^{l-1})$.

Proof. The argument is very similar to that used in the proof of Lemma 3, I. If $d = p^tD$ where $(p, D) = 1$, then $p^{l*} = p^{t+l}$ so that $N(p^l)$ is the number of solutions of

$$\sum_{\nu=1}^{s} P(x_\nu) \equiv n \pmod{p^l}, \quad 0 \leqq x_\nu < p^{t+l}, \quad p \nmid \text{every } P_0(x_\nu). \tag{2.21}$$

Hence $N(p^l)$ is equal to D^{-s} times the number of solutions of

$$\sum_{\nu=1}^{s} P(x_\nu) \equiv n \pmod{p^l}, \quad 0 \leqq x_\nu < Dp^{t+l}, \quad p \nmid \text{every } P_0(x_\nu). \tag{2.22}$$

For every x_ν in (2.22) let $z_\nu = [x_\nu/(Dp^{t+l-\theta-1})]$ so that

$$x_\nu = y_\nu + z_\nu Dp^{t+l-\theta-1}, \qquad \begin{array}{l} 0 \leqq y_\nu < Dp^{t+l-\theta-1}, \\ 0 \leqq z_\nu < p^{\theta+1}. \end{array}$$

Then by Lemma 1 we may write (2.22) in the form

$$\begin{gathered} \sum_{\nu=1}^{s} P(y_\nu) + p^{l-1} \sum_{\nu=1}^{s} z_\nu P_0(y_\nu) \equiv n \pmod{p^l}, \\ 0 \leqq y_\nu < Dp^{t+l-\theta-1}, \; 0 \leqq z_\nu < p^{\theta+1}, \; p \nmid \text{every } P_0(y_\nu). \end{gathered}$$

It follows that to each solution of (2.22) there corresponds a solution of the two congruences

$$\sum_{\nu=1}^{s} P(y_\nu) \equiv n \pmod{p^{l-1}}, \quad 0 \leqq y_\nu < Dp^{t+l-\theta-1}, \quad p \nmid \text{every } P_0(y_\nu), \tag{2.23}$$

$$\sum_{\nu=1}^{s} z_\nu P_0(y_\nu) \equiv p^{-l+1}\Big(n - \sum_{\nu=1}^{s} P(y_\nu)\Big) \pmod{p}, \quad 0 \leqq z_\nu < p^{\theta+1}. \tag{2.24}$$

By the same method of proof as that used in Lemma 3, I, it can be shown that (2.23) has $D^s p^{-\theta s} N(p^{l-1})$ solutions and that (2.24) has $p^{\theta s+s-1}$ solutions. Hence (2.22) has $p^{\theta s+s-1}D^s p^{-\theta s}N(p^{l-1})$ solutions. Then

$$N(p^l) = D^{-s}p^{\theta s+s-1}D^s p^{\theta s}N(p^{l-1}) = p^{s-1}N(p^{l-1}),$$

and this completes the proof.

3. The proof of Theorem 1 when $a \equiv 4c \pmod 8$. As explained in I, Theorem 1 is a consequence of Theorem 3, I, and this theorem in turn depends upon the fact that

$$(3.11) \qquad p^{-l(s-1)}M(p^l) \geqq p^{-\gamma(s-1)}, \qquad l \geqq \gamma, \qquad s \geqq 9,$$

where γ is some fixed integer. If the proof of (3.11) as presented in I is examined, it is found that the restriction $a \not\equiv 4c \pmod 8$ was used only when $b \equiv 6c \pmod 8$, c odd. This was in (6.41). Hence in this section we shall prove that (3.11) is true when $P(x)$ has the form (1.1) with $a \equiv 4c$, $b \equiv 6c \pmod 8$ and c odd.

In sections 4 and 5 of I we have shown that the number of solutions of [3]

$$(3.12) \qquad \sum_{\nu=1}^{s} P(vx_\nu + t) \equiv n \pmod{p^l}, \qquad 0 \leqq x_\nu < p^l$$

is $\geqq p^{(l-2)(s-1)}$ when $p \geqq 3$. Since $p^{l*} = p^l$ when $P(x)$ is a cubic polynomial and $p > 3$, it is evident that the number of solutions of (3.12) is the same as the number of solutions of

$$(3.13) \qquad \sum_{\nu=1}^{s} P(y_\nu) \equiv n \pmod{p^l}, \qquad 0 \leqq y_\nu < p^{l*}.$$

The congruence (3.13) has $M(p^l)$ solutions by definition. Hence we have $M(p^l) \geqq p^{(l-2)(s-1)}$ when $p > 3$, and this proves (3.11) when $p > 3$. In a similar manner it can be shown that (3.11) is true with $\gamma = 2$ when $p = 3$ and $3 \mid a$, for in this case $v = 1$.

Thus two cases, $p = 3$, $3 \nmid a$ and $p = 2$, remain. We shall dispose of them in Lemmas 3 and 4.

LEMMA 3. *If $p = 3$, $3 \nmid a$ then (3.11) is true with $\gamma = 1$.*

Proof. In this case we have $d = 3$, $3^* = 9$, $\Phi(x) = 3P(x)$,

$$\Phi'(x) = 3a(x^2 + x)/2 + (6b - 3a)x/2 - (a + 3b - 6c)/2.$$

From this equation it is evident that $3 \nmid \Phi'(x)$ for all values of x since $3 \nmid a$. Hence $\theta = 0$ and $P_0(x) = \Phi'(x)$.

We distinguish two cases. 1). Suppose $3 \mid b$. Since

$$P(x+1) + 2P(9-x) + P(x-1) \equiv ax \pmod 3$$

there exists a solution x_0 of the congruence

$$P(x_0+1) + 2P(9-x_0) + P(x_0-1) \equiv n \pmod 3.$$

Hence the congruence

$$\sum_{\nu=1}^{5} P(x_\nu) \equiv n \pmod 3, \qquad 0 \leqq x_\nu < 3^*$$

[3] See (1.41)-(1.44) of I for the definition of v and t.

has the solution $x_1 = 0$, $x_2 = x_0 + 1$, $x_3 = x_4 = 9 - x_0$, $x_5 = x_0 - 1$. Also, $P_0(0) = -(a + 3b - 6c)/2$ and this expression is not divisible by 3. Hence for $s \geqq 5$, $l \geqq 2 = \max(2\theta + 2 - t, \theta + 2)$, it follows from Lemma 2 that

$$(3.2) \quad M(3^l) \geqq N(3^l) = 3^{s-1}N(3^{l-1}) = \cdots = 3^{(l-1)(s-1)}N(3) \geqq 3^{(l-1)(s-1)}.$$

This proves the lemma when $3 \mid b$.

2). Suppose $3 \nmid b$. In this case we have

$$P(x) + P(9 - x) \equiv bx^2 \pmod 3$$

and so there is a solution of the congruence [4]

$$\sum_{\nu=1}^{3} [P(x_\nu) + P(9 - x_\nu)] \equiv n \pmod 3.$$

As before we have a solution $y_1 = 0$, $y_{2\nu} = x_\nu$, $y_{2\nu+1} = 9 - x_\nu$, $\nu = 1, 2, 3$, of the congruence

$$\sum_{\mu=1}^{7} P(y_\mu) \equiv n \pmod 3, \qquad 0 \leqq y_\mu < 3^*,$$

in which $P_0(0)$ is not divisible by 3. Thus (3.2) is proved in this case also with $s \geqq 7$.

LEMMA 4. *If $p = 2$ then (3.11) is true with $\gamma = 5$.*

Proof. In this case we have $d = 1$ or 3,

$$\Phi(x) = da(x^3 - x)/6 + db(x^2 - x)/2 + dcx$$
$$\Phi'(x) = da(x^2 + x)/2 + d(2b - a)x/2 - d(a + 3b - 6c)/6.$$

Using the relations $a \equiv 4c$, $b \equiv 6c \pmod 8$ it is easily seen that $\Phi'(x) \equiv 0 \pmod 4$ for all values of x, but that $\Phi'(x) \not\equiv 0 \pmod 8$ for all values of x. Hence $\theta = 2$.

If n is even let $x_0 = 0$ or 1 according as $n \equiv 2$ or $n \equiv 0 \pmod 4$. Then since $P(0) = 0$ and $P(1) = c$ is odd, we have $n - 2P(x_0) \equiv 2 \pmod 4$. If n is odd let $r = 1$ or 3 according as $n \equiv 3c$ or $n \equiv c \pmod 4$. Then we have $n - rP(1) \equiv 2 \pmod 4$. This shows that we can always write n in the form

$$n = rP(x_0) + 2n_1,$$

where $r = 1$, 2, or 3, and n_1 is odd.

[4] E. Landau, *Vorlesungen über Zahlentheorie*, Bd. I, Theorem 301.

Now $P(x) + P(32 - x) \equiv bx^2 \pmod{32}$ and since $b/2$ is odd there exists a solution of the congruence [5]

$$(b/2) \sum_{\nu=1}^{3} x_\nu^2 \equiv n_1 \pmod{16}$$

in which x_1 is odd. It follows that

$$\sum_{\nu=1}^{3} [P(x_\nu) + P(32 - x_\nu)] \equiv 2n_1 \pmod{32}, \qquad 0 \leqq x_\nu < 16$$

and hence that

$$rP(x_0) + \sum_{\nu=1}^{3} [P(x_\nu) + P(32 - x_\nu)] \equiv n \pmod{32}.$$

Moreover at least one of $P_0(x_1)$ and $P_0(32 - x_1)$ is not divisible by 2. For if both were divisible by 2 we should have

$$dbx_1/2 \equiv P_0(x_1) - P_0(32 - x_1) \equiv 0 \pmod{2},$$

whereas d, $b/2$, and x_1 are all odd. Then for

$$s \geqq 9 \geqq r + 6, \qquad l \geqq 6 = \max(2\theta + 2 - t, \theta + 2),$$

we have

$$M(2^l) \geqq N(2^l) = 2^{s-1}N(2^{l-1}) = \cdots = 2^{(l-5)(s-1)}N(32) \geqq 2^{(l-5)(s-1)}.$$

This completes the proof for the case $p = 2$.

THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIFORNIA.

[5] E. Landau, *Vorlesungen über Zahlentheorie*, Bd. I, Theorem 301.

THE CONJUNCTIVE EQUIVALENCE OF PENCILS OF HERMITIAN AND ANTI-HERMITIAN MATRICES.*

By John Williamson.

Let K be a commutative field of characteristic zero and let $K(i)$ be a quadratic adjunction field of K, where i is a zero of the polynomial $x^2 - a$, irreducible in K. If A is a matrix with elements in $K(i)$, or more shortly a matrix over $K(i)$, the matrix A^* is defined to be the conjugate transposed of A, so that $A^* = \bar{A}'$. In particular, if A is a matrix over K, $A^* = A'$, the transposed of A. Let

(1) $$\Lambda = rA + sB,$$

be a pencil of matrices, in which

(2) $$A^* = \epsilon A, \qquad B^* = \delta B, \qquad \epsilon, \delta = \pm 1,$$

so that A is either hermitian or anti-hermitian and so is B. Let $\Lambda_1 = rA_1 + sB_1$ be another such pencil. Then the two pencils Λ and Λ_1 are said to be *conjunctively equivalent*, if there exists a non-singular matrix P over $K(i)$ such that

$$P\Lambda P^* = \Lambda_1;$$

that is, if $PAP^* = A_1$ and $PBP^* = B_1$. When the matrices A, B, A_1, B_1 are all matrices over K, the two pencils are said to be *congruently equivalent*, if there exists a non-singular matrix P over K, such that

$$P\Lambda P' = \Lambda_1.$$

There are accordingly two distinct problems to be considered; (a) to determine necessary and sufficient conditions that two pencils, which satisfy (2), be conjunctively equivalent and (b) to determine necessary and sufficient conditions that two such pencils be congruently equivalent. Problem (a) has been solved completely, for the case in which $\epsilon = \delta = 1$, when K is the field of all real numbers [1] and also, under the restriction that B be non-singular, when K is a commutative field of characteristic zero.[2] In the other two cases,

* Received November 30, 1936.

[1] H. W. Turnbull, "On the equivalence of pencils of hermitian forms," *Proceedings of the London Mathematical Society*, vol. 39 (1935), pp. 232-248; M. H. Ingraham and K. W. Wegner, "The equivalence of pairs of hermitian matrices," *Transactions of the American Mathematical Society*, vol. 38 (1935), pp. 145-162. In both papers a treatment of singular pencils is given.

[2] John Williamson, "The equivalence of non-singular pencils of matrices in an

in which one or both of ϵ, δ have the value -1, problem (a) may be reduced to the case in which $\epsilon = \delta = 1$. For example, let $\epsilon = -1$, $\delta = 1$, so that $A = -A^*$, $B = B^*$. Then, $(iA)^* = iA$ and

$$(3) \qquad \Lambda = tC + sB,$$

where $t = r/i$, $C = C^*$, $B = B^*$ and (3) is a pencil of hermitian matrices. Hence, when Λ is a non-singular pencil, problem (a) is completely solved.

Problem (b) has been solved for the case, in which $\epsilon = \delta = 1$, when K is the real field [3] and also, under the restriction that B be non-singular, when K is a general commutative field of characteristic zero, (I). It is however not possible to reduce the other cases to this one and they must be considered separately. The problem has also been solved for a non-singular pencil and general field K, when $\epsilon = 1$, $\delta = -1$.[4] The remaining problem when $\epsilon = \delta = -1$, so that both A and B are skew symmetric matrices, is considered here. It is shown that two skew symmetric matrices are congruently equivalent, if they have the same kronecker minimal indices and the same invariant factors.[5] This is a much simpler result than those obtained in the other cases; but it is only natural that this be so, since two skew symmetric matrices of the same rank are congruently equivalent, while the same is not true of two symmetric matrices.

Since, when A and B are both skew symmetric matrices of odd order, every matrix of the pencil Λ is singular, a treatment of singular pencils is absolutely necessary. Accordingly the conjunctive or congruent equivalence of two general pencils Λ, which satisfy (2), is first considered, so that the solution of both problems (a) and (b) is completed in all cases. The method here adopted is quite distinct from that used in the discussion of non-singular pencils of hermitian matrices,[6] and has the advanatge that at no stage is a change made in the basis of the pencil.

Section 1 is devoted to the proofs of subsidiary lemmas, section 2 to the consideration of singular pencils, section 3 to that of non-singular pencils in which B is singular, and 4 to the reduction of a non-singular pencil of skew symmetric matrices to canonical form.

---

arbitrary field," *American Journal of Mathematics*, vol. 57 (1935), pp. 475-490. This paper will be referred to as I.

[3] Turnbull, *loc. cit.*

[4] John Williamson, "On the algebraic problem concerning the normal forms of linear dynamical systems," *American Journal of Mathematics*, vol. 58 (1936), pp. 141-163. This paper will be referred to as II.

[5] This is a well known result in case K is algebraically closed. See L. E. Dickson, *Modern Algebraic Theories*, p. 125, or C. C. MacDuffee, *The Theory of Matrices*, p. 61.

[6] Turnbull, *loc. cit.*; Wegner and Ingraham, *loc. cit.*

1. If $\Lambda = rA + sB$, we define the matrix pencil Λ'' by

$$\Lambda'' = \epsilon rA^* + \delta sB^*. \tag{4}$$

Let ρ_j be the matrix pencil

$$\rho_j = \begin{pmatrix} r & s & 0 & \cdot & 0 & 0 \\ 0 & r & s & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & s & 0 \\ 0 & 0 & 0 & \cdot & r & s \end{pmatrix}, \tag{5}$$

of j rows and $j+1$ columns. Then ρ_j'' is a matrix pencil of $j+1$ rows and j columns, while the matrix pencil

$$R_j = \begin{pmatrix} \rho_j & 0 \\ 0 & \rho_j'' \end{pmatrix} \tag{6}$$

is a square matrix of $2j+1$ rows and columns. Let

$$N_j = rE_j + sU_j, \tag{7}$$

where E_j and U_j are respectively the unit matrix and the auxiliary unit matrix of order j.[7] We now prove a few elementary lemmas involving the matrices R_j and N_j.

LEMMA I. *Let S and T be two matrices, which satisfy*

$$SR_q = R_p''T. \tag{8}$$

Then, if $p < q$, the first row of S is zero; if $p = q$,

$$S = \begin{pmatrix} 0 & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}, \tag{9}$$

where σ_{12} and σ_{21} are diagonal matrices of orders $p+1$ and p respectively.

Proof. Since S and T satisfy (8) we may write S and T as two rowed square matrices of matrices,

$$S = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}, \qquad T = \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{pmatrix},$$

where σ_{11} is a matrix of $p+1$ rows and q columns, σ_{12} a matrix of $p+1$ rows and $q+1$ columns etc. From (8) we immediately deduce the four equations

$$\sigma_{11}\rho_q = \rho_p''\tau_{11}, \quad \sigma_{12}\rho_q'' = \rho_p''\tau_{12}, \quad \sigma_{21}\rho_q = \rho_p\tau_{21}, \quad \sigma_{22}\rho_q'' = \rho_p\tau_{22}. \tag{10}$$

[7] See Turnbull and Aitken, *Canonical Matrices*, p. 62.

12

The first of these equations is of the form,

$$C\rho_q = \rho_p''D, \tag{11}$$

where $C = (c_{ij})$ is a matrix of $p+1$ rows and q columns, while $D = (d_{ij})$ is a matrix of p rows and $q+1$ columns. As a consequence of (11) we have

$$rc_{ij} + sc_{i,j-1} = \epsilon rd_{ij} + \delta sd_{i-1,j}, \quad (i = 1, 2, \cdots, p+1,\ j = 1, 2, \cdots, q+1),$$

with the understanding that

$$c_{i,0} = d_{0,j} = c_{i,q+1} = d_{p+1,j} = 0. \tag{12}$$

Therefore,

$$c_{ij} = \epsilon d_{ij}, \quad c_{i,j-1} = \delta d_{i-1,j}, \quad c_{ij} = \epsilon\delta c_{i+1,j-1}, \tag{13}$$

and we see from the last of these equations that the elements in any counter diagonal of C are the same except perhaps for sign. But, as a consequence of (12) and (13),

$$c_{p+1,j} = \epsilon d_{p+1,j} = 0$$

and also

$$c_{i-1,1} = \epsilon\delta c_{i,0} = 0, \qquad (i = 2, 3, \cdots, p+1).$$

Hence the first column and the last row of C are zero, and consequently C is zero. Therefore σ_{11} is zero.

The second of equations (10) is of the form $C\rho_q'' = \rho_p''D$, where $C = (c_{ij})$ is a matrix of $p+1$ rows and $q+1$ columns, while $D = (d_{ij})$ is a matrix of p rows and q columns. Consequently

$$c_{ij} = \epsilon d_{ij}, \quad c_{i,j+1} = \delta d_{i-1,j}, \quad c_{ij} = \epsilon\delta c_{i+1,j+1},$$
$$(i = 1, 2, \cdots, p+1;\ j = 1, 2, \cdots, q),$$

so that the elements in any diagonal of C differ at most in sign. But

$$c_{1,j+1} = \delta d_{0,j} = 0 \quad \text{and} \quad c_{p+1,j} = \epsilon d_{p+1,j} = 0.$$

Hence

$$c_{1j} = 0, \qquad (j = 2, \cdots, q+1),$$
$$c_{p+1,j} = 0, \qquad (j = 1, 2, \cdots, q).$$

Therefore, if the number of rows of C is less than the number of columns, i. e. if $p < q$, $C = 0$ and, if $p = q$, C is a diagonal matrix. Hence, if $p < q$, $\sigma_{12} = 0$ and, if $p = q$, σ_{12} is a diagonal matrix. Moreover, if $p = q$, D is also a diagonal matrix, whose first element is ϵ times the first element of C. We have as a result the corollary:

COROLLARY I. *If $T = S^*$ and σ_{12} is non-singular, σ_{21} is non-singular.*

The proofs of the following lemmas, which are similar to the above, are omitted.

LEMMA 2. *If* $SN_i = R_q''T$, *the first row of* S *is zero.*

LEMMA 3. *If* M *is a square matrix of order* t *and* $S(rM + sE_t) = R_q''T$ *the first row of* S *is zero.*

LEMMA 4. *If* $SN_i = N_j''T$ *and* $i < j$, *the first row of* S *is zero.*

LEMMA 5. *If* $S(rM + sE) = N_i''T$, *the first row of* S *is zero.*

When G is a square matrix of order n, we may consider G as a matrix of matrices and write

$$G = (G_{ij}), \qquad (i, j = 1, 2, \cdots, t),$$

where G_{ij} is a matrix of n_i rows and n_j columns. If H is a second n-rowed square matrix

$$H = (H_{ij}), \qquad (i, j = 1, 2, \cdots, t),$$

where H_{ij} is also a matrix of n_i rows and n_j columns we shall say that G and H are *similarly partitioned.* If $G_{ij} = 0$, when $i \neq j$, we shall call G a *diagonal block* matrix and write

$$G = [G_{11}, G_{22}, \cdots, G_{tt}].$$

2. Let Λ be the pencil defined by (1) and (2) so that

$$\Lambda = \Lambda''. \tag{14}$$

Let Λ annihilate the column vector u of dimension n, whose elements are polynomials in r and s with coefficients in $K(i)$. Then

$$\Lambda u = 0, \text{ identically in } r \text{ and } s,$$

and consequently

$$0 = u''\Lambda'' = u''\Lambda, \text{ by (14)}.$$

Since u'' is the row vector obtained from u^* by replacing r by ϵr and s by δs, the degree of u'' in r and s is the same as the degree of u. Therefore *the set of kronecker minimal row indices* [8] *coincides with the set of minimal column indices.* In particular, if $u_1, u_2, \cdots, u_p$ form a complete set of linearly independent vectors over $K(i)$ annihilated by Λ and, if V is any non-singular matrix, whose first p columns are the vectors $u_1, u_2, \cdots, u_p$, then

$$V^*\Lambda V = [0, \Lambda_1], \tag{15}$$

[8] Turnbull and Aitken, *op. cit.*, pp. 119-125.

where Λ_1 is a pencil of the same type as Λ but of order $n - p$. Moreover none of the minimal row (or column indices) of Λ_1 is zero. Since (15) is a conjunctive transformation and since, when A and B are matrices over K, $V^* = V'$, we may consider the pencil Λ_1 instead of Λ. Accordingly without any loss of generality we may assume that *no minimal row or column index of* Λ *is zero.*

Let Π be any pencil equivalent to Λ in the more general sense that there exist two non-singular matrices Q and P, such that

$$Q\Lambda P = \Pi.$$

Then,

$$P^*\Lambda P = P^*Q^{-1}\Pi = H\Pi,$$

so that the pencil $H\Pi$ is conjunctively equivalent to Λ. Accordingly, as a consequence of (14),

$$H\Pi = \Pi'' H^*. \tag{16}$$

We now prove

LEMMA 6. *Let* $\Pi = [\Pi_1, \Pi_2]$ *be a diagonal block matrix, where* Π_i *is of order* n_i, *and let the matrix* H_1, *formed from the first* n_1 *rows and columns of* H, *be non-singular. Then, if equation* (16) *is satisfied, there exists a non-singular matrix* W *such that*

$$W^*H\Pi W = [H_1\Pi_1, H_2\Pi_2],$$

is also a diagonal block matrix.

Proof. Let $H = (H_{ij})$, $i, j = 1, 2$, be a partition of H similar to that of Π. Then $H_{11} = H_1$ is non-singular.

Equation (16) implies

$$H_{ij}\Pi_j = \Pi_i'' H^*_{ji}, \qquad (i, j = 1, 2),$$

and consequently

$$\Pi_1 (H_{11}^{-1})^* H^*_{21} = H_{11}^{-1}\Pi_1'' H^*_{21} = H_{11}^{-1} H_{12}\Pi_2.$$

It now follows by a simple calculation that, if

$$W = \begin{pmatrix} E_1 & 0 \\ -H_{21}H_{11}^{-1} & E_2 \end{pmatrix},$$

where E_i is the unit matrix of order n_i,

$$W\Pi W^* = [H_1, H_2][\Pi_1, \Pi_2].$$

Since the matrix W depends solely on H we have the

COROLLARY. *If H is a matrix over K, W is a matrix over K and $W^* = W'$.*

We now take Π in the canonical form [9]

$$\Pi = [R_{j_1}, R_{j_2}, \cdots, R_{j_k}, N_{t_1}, N_{t_2}, \cdots, N_{t_m}, rM + sE]. \tag{17}$$

In (17) each matrix R_{j_i} is defined by (6) and $j_1, j_2, \cdots, j_k$ form a complete set of minimal row (or column) indices, each matrix N_{t_i} is defined by (7) and corresponds to an elementary factor r^{t_i} of Λ; and $rM + sE$ is a pencil whose second member is non-singular and can therefore be taken in this simplified form.

For convenience we write (17) in the form

$$\Pi = [\pi_1, \pi_2, \cdots, \pi_{k+m+1}], \tag{18}$$

where $\pi_i = R_{j_i}$, $i < k$; $\pi_{k+i} = N_{t_i}$, $i < m$; $\pi_{k+m+1} = rM + sE$.

If $H = (H_{pq})$, $p, q = 1, 2, \cdots, k + m + 1$, is a partition of H similar to that of Π in (18), it follows from (16) that

$$H_{pq}\pi_q = \pi_p'' H^*_{qp}. \tag{19}$$

Let the integers $j_1, j_2, \cdots, j_k$ in (17) be so arranged that

$$j_1 = j_2 = \quad = j_c < j_{c+1} \leqq \cdots \leqq j_k.$$

If H_{11} is singular and, for some value of $p \leqq c$, H_{pp} is non-singular, by a conjunctive transformation involving an interchange of rows and the same interchange of columns, we may interchange H_{11} and H_{pp} without disturbing Π. By (19) and Lemmas (1), (2) and (3), the first row of H_{pq} is zero, when $p \leqq c$ and $q > c$, and contains at most one element h_{pq}, in the $(j_1 + 1)$-th place, different from zero when $p \leqq c$ and $q \leqq c$. Further, by the corollary to Lemma (1), h_{pp} is zero, if, and only if, H_{pp} is singular. Therefore, if H_{11} is singular, $h_{11} = 0$, and since H is non-singular at least one element in the first row of H is different from zero. Accordingly, for at least one value of $p \leqq c$, $h_{1p} \neq 0$. Without any loss of generality, then, we may assume that the first row of H_{12} contains the element h_{12} distinct from zero.

Let E be the unit matrix of order $2j + 1 = 2j_1 + 1$ and let

$$W_1 = \begin{pmatrix} E & E \\ 0 & E \end{pmatrix}, \qquad W_2 = \begin{pmatrix} E & iE \\ 0 & E \end{pmatrix}. \tag{20}$$

---

[9] W. Ledermann, "Reduction of singular pencils of matrices," *Proceedings of the Edinburgh Mathematical Society*, vol. 4 (1935), series 2, pp. 92-105.

Then

$$W_1(H_{pq})[R_j, R_j]W^*_1 = (G_{pq})[R_j, R_j],$$
$$W_2(H_{pq})[R_j, R_j]W^*_2 = (F_{pq})[R_j, R_j], \qquad (p, q = 1, 2),$$

where

$$G_{11} = H_{11} + H_{12} + H_{21} + H_{22}, \qquad F_{11} = H_{11} + iH_{21} - iH_{12} - i^2H_{22}.$$

Let g_{11} and f_{11} be the elements in the first row and $(j+1)$-th column of the matrices G_{11} and F_{11} respectively. Then

$$(21) \quad g_{11} = h_{12} + h_{21} + h_{11} + h_{22}, \quad f_{11} = i(h_{21} - h_{12}) + h_{11} - i^2h_{22}.$$

If H_{11} and H_{22} are both singular, $h_{11} = h_{22} = 0$ and at least one of g_{11} or f_{11} is non-zero, as otherwise h_{12} would be zero. Therefore at least one of G_{11} or F_{11} is non-singular.

If, however, we are restricted to congruent transformations over the field K, we are not at liberty to use the transformation W_2. But, if

$$Q = [-E_{j+1}, E_j, E_{2j+1}],$$
$$[R_j, R_j]Q = [-E_j, E_{j+1}, E_{2j+1}][R_j, R_j],$$

and

$$Q(H_{pq})[R_j, R_j]Q' = (K_{pq})[R_j, R_j], \qquad (p, q = 1, 2),$$

where

$$K_{12} = [-E_{j+1}, E_j]H_{12}, \qquad K_{21} = H_{21}[-E_j, E_{j+1}].$$

Accordingly, if $W_2 = W_1Q$, equation (21) becomes

$$g_{11} = h_{12} + h_{21}, \qquad f_{11} = h_{21} - h_{12},$$

so that once again either G_{11} or F_{11} is non-singular. Therefore, we may suppose that the necessary transformation has already been made and that H_{11} is non-singular. As a consequence of Lemma (6), there exists a non-singular matrix P_1, such that

$$P_1H\Pi P^*_1 = [H_{11}R_{j_1}, H_2\Pi_2].$$

By repeating this process k times we prove the existence of a non-singular matrix P such that

$$(22) \qquad PH\Pi P^* = [H_1R_{j_1}, H_2R_{j_2}, \cdots, H_kR_{j_k}, H_{k+1}\Pi_{k+1}],$$

where $\Pi_{k+1} = [N_{i_1}, N_{i_2}, \cdots, N_{i_t}, rM + sE]$. All matrices H_i in (22) are non-singular and satisfy the equations

$$(23) \quad H_iR_{j_i} = R_{j_i}''H^*_i, \ (i = 1, 2, \cdots, k), \qquad H_{k+1}\Pi_{k+1} = \Pi''_{k+1}H^*_{k+1}.$$

To simplify equations (23) we write $H_i = S$, $R_{j_1} = R_p$, and with the notation of Lemma 1, have

$$S = \begin{pmatrix} 0 & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix},$$

where σ_{12} and σ_{21} are non-singular and satisfy

$$\sigma_{12}\rho_p'' = \rho_p''\sigma^*_{21}, \qquad \sigma_{22}\rho_p'' = \rho_p\sigma^*_{22}. \tag{24}$$

Let W be the matrix

$$W = \begin{pmatrix} \sigma_{12}^{-1} & 0 \\ -\frac{1}{2}\sigma_{22}\sigma_{12}^{-1} & E_p \end{pmatrix}.$$

Then,

$$R_pW^* = \begin{pmatrix} \rho_p(\sigma_{12}^{-1})^* & -\frac{1}{2}\rho_p(\sigma_{12}^{-1})^*\sigma^*_{22} \\ 0 & \rho_p'' \end{pmatrix} = \begin{pmatrix} \sigma_{21}^{-1}\rho_p & -\frac{1}{2}\sigma_{21}^{-1}\sigma_{22}\rho_p'' \\ 0 & \rho_p'' \end{pmatrix} \text{ by (24)}$$

$$= \begin{pmatrix} \sigma_{21}^{-1} & -\frac{1}{2}\sigma_{21}^{-1}\sigma_{22} \\ 0 & E_{p+1} \end{pmatrix} R_p,$$

and

$$WSR_pW^* = \begin{pmatrix} 0 & E_{p+1} \\ E_p & 0 \end{pmatrix} R_p.$$

Therefore each matrix H_i in (22) may be reduced to the form I_{j_i}, where

$$I_j = \begin{pmatrix} 0 & E_{j+1} \\ E_j & 0 \end{pmatrix}. \tag{25}$$

Accordingly we have proved,

THEOREM I. *There exists a non-singular matrix P such that*

$$P\Lambda P^* = [I_{j_1}R_{j_1}, I_{j_2}R_{j_2}, \cdots, I_{j_k}R_{j_k}, H_{k+1}\Pi_{k+1}], \tag{26}$$

where I_j is defined by (25) and R_j by (6), while the pencil $H_{k+1}\Pi_{k+1}$ is non-singular.

COROLLARY 1. *If the elements of A and B lie in K, the matrix P lies in K and $P^* = P'$.*

COROLLARY 2. *The canonical form on the right of (26) is determined completely by the minimal indices of Λ and the non-singular core $H_{k+1}\Pi_{k+1}$.*

3. In considering the reduction of the non-singular case we could, by a change of the basis of the pencil, reduce it to one in which the second member is non-singular. As a change of basis is quite distinct in nature from the conjunctive transformations of the pencil it is more satisfactory to proceed as follows.[10] For simplicity we write

$$\Pi_{k+1} = \Pi = [N_{t_1}, N_{t_2}, \cdots, N_{t_m}, rM + sE], \tag{27}$$

and

$$H_{k+1} = H = (H_{ij}), \qquad (i, j = 1, 2, \cdots, m+1), \tag{28}$$

[10] Cf. Ledermann, *loc. cit.*

where (28) is a partition of H similar to that of Π in (27). If

$$(29)\qquad t_1 = t_2 = \quad = t_c > t_{c+1} \geqq \cdots \geqq t_m,$$

it follows from Lemmas (4) and (5) that the first row of $H_{1p} = 0$, when $p > c$, and that at most one element in the first row of H_{1q} is different from zero, when $p \leqq c$. If H_{11} is singular, we may suppose, as in the previous section, that H_{12} is non-singular and, since $H_{12} = \epsilon H^*_{21}$, that $\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}$ is non-singular. Therefore by several applications of Lemma 6 we may reduce the pencil $H\Pi$ to a diagonal block form where each block is either of type (α) $H_{11}N_t$ or of type (β) $\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}\begin{pmatrix} N_t & 0 \\ 0 & N_t \end{pmatrix}$.

It is now necessary to consider three distinct cases;

(1) $\epsilon = \delta = 1$, (2) $\epsilon = +1$, $\delta = -1$, (3) $\epsilon = \delta = -1$.

Case 1. Both matrices in the pencil are hermitian or both symmetric. A block of type (β) may be reduced to two of type (α) (I page 481). The matrix H_{11} in (α) is of the form,

$$H_{11} = T_t(g_1E_t + g_2U_t + \cdots + g_{t-1}U_t^{t-1}),$$

where T_t is the counter unit matrix of order t.

Finally H_{11} may be reduced to the form

$$(30)\qquad H_{11} = g_1T_t,$$

where g_1 lies in K.

In (30) g_1 is not uniquely determined. In fact, the diagonal matrix $G = [g_1, g_2, \cdots, g_c]$, where c is defined by (29), may be replaced by the diagonal matrix $F = [f_1, f_2, \cdots, f_c]$, provided F and G are conjunctively equivalent (I pages 482-487), or

THEOREM 2. *If r^t occur exactly c times among the elementary factors of a pencil of hermitian matrices Λ, in the canonical form for Λ there is a block $[g_1T_t, g_2T_t, \cdots, g_cT_t]$. The diagonal matrix $[g_1, g_2, \cdots, g_c]$ is determined apart from a conjunctive transformation.*

Theorems 1 and 2 together with the theorem in I, page 487, give a complete solution of problem (a).

Case 2. As stated in the introduction, it is only necessary to consider congruent transformations of pencils with elements in K. The matrix H_{11} is symmetric, while the matrix $H_{11}U_t$ is skew symmetric. Therefore

$$(31)\qquad H_{11}U = -U'H'_{11} = -U'H_{11}, \text{ where } U = U_t.$$

Let

$$X = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & 0 & 1 \\ 0 & \cdot & \cdot & \cdot & -1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (-1)^{t-1} & \cdot & \cdot & \cdot & 0 & 0 \end{pmatrix}.$$

Then $XU = -U'X$, and as a consequence of (31) $H_{11} = XG$, where G is commutative with U. Hence

$$H_{11} = Xf(U), \tag{32}$$

where $f(U)$ is a polynomial in U, with coefficients in K.

Since $H_{11} = H'_{11}$,

$$Xf(U) = f(U')X' = X'f(-U).$$

Hence, if t is odd,

$$f(U) = f(-U) = g(U^2),$$

while, if t is even,

$$f(U) = -f(-U) = Ug(U^2).$$

Let t be odd. The congruent transformation by the matrix W_1 of (20) reduces a block of type (β) to one in which H_{11} is non-singular. Therefore we need only consider blocks of type (α). Let the matrix $f(U)$ in (32) be written as

$$f(u) = gE + \gamma, \text{ where } \gamma = \gamma(U^2) = U^2\gamma_1(U^2).$$

Then, if $W = E - \frac{1}{2}g^{-1}\gamma$,

$$\begin{aligned} W'Xf(U)NW &= X(E - \tfrac{1}{2}\gamma g^{-1})(gE + \gamma)(E - \tfrac{1}{2}g^{-1}\gamma)N, \\ &= X(gE + \gamma^2\phi)N, \end{aligned}$$

where ϕ is a polynomial in U^2. Since γ^2 contains a factor U^4, it is possible by a succession of such transformations to reduce H_{11} to the form XgE. The element g is not unique. By an argument similar to that of II, pages 152-154, we have

THEOREM 3. *If t is odd and r^t occurs c times among the elementary factors of the pencil Λ, in the canonical form of Λ occurs a block matrix $[g_1X_tN_t, g_2X_tN_t, \cdots, g_cX_tN_t]$, where the diagonal matrix $[g_1, g_2, \cdots, g_c]$ is determined apart from a congruent transformation.*

Let t be even. Then H_{11} is singular and only blocks of type (β) may occur. It is easily shown that

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \begin{pmatrix} XUf_{11}(U) & Xf_{12}(U) \\ Xf_{21}(U) & XUf_{22}(U) \end{pmatrix} = \gamma + \phi,$$

where $\gamma = \begin{pmatrix} 0 & Xf_{12}(U) \\ Xf_{21}(U) & 0 \end{pmatrix}$, $\gamma[N, N] = [N'', N'']\gamma'$ and

$$\phi[N, N] = [N'', N'']\phi'.$$

Therefore,

$$\begin{aligned}(E - \tfrac{1}{2}\phi\gamma^{-1})(\gamma + \phi)[N, N](E - \tfrac{1}{2}\phi\gamma^{-1})' &= (E - \tfrac{1}{2}\phi\gamma^{-1})(\gamma + \phi)\,(E - \tfrac{1}{2}\gamma^{-1}\phi)[N, \\ &= (\gamma - \phi\gamma^{-1}\phi + \tfrac{1}{4}\phi\gamma^{-1}\phi\gamma^{-1}\phi)[N, N], \\ &= (\gamma_1 + \phi_1)[N, N],\end{aligned}$$

where

$$\gamma_1 = \begin{pmatrix} 0 & Xg_{12}(U) \\ Xg_{21}(U) & 0 \end{pmatrix} \quad \text{and} \quad \phi_1 = \begin{pmatrix} XU^3g_{11}(U) & 0 \\ 0 & XU^3g_{22}(U) \end{pmatrix}.$$

Since γ is non-singular, γ_1 is non-singular, and we may repeat this process with γ replaced by γ_1 and ϕ by ϕ_1. After $t/2$ repetitions the matrix $\phi_{t/2}$ is the zero matrix, since $U^t = 0$. Hence a block of type (β) may be reduced to the form

$$(33) \qquad \begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix}\begin{pmatrix} N & 0 \\ 0 & N \end{pmatrix},$$

where $H_{12}N = N''H'_{21}$. But,

$$(34) \qquad \begin{aligned}&\begin{pmatrix} E & 0 \\ 0 & H_{21}^{-1} \end{pmatrix}\begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix}\begin{pmatrix} N & 0 \\ 0 & N \end{pmatrix}\begin{pmatrix} E & 0 \\ 0 & H_{21}^{-1} \end{pmatrix}' \\ &= \begin{pmatrix} 0 & H_{12} \\ E & 0 \end{pmatrix}\begin{pmatrix} E & 0 \\ 0 & H_{12}^{-1} \end{pmatrix}\begin{pmatrix} N & 0 \\ 0 & N'' \end{pmatrix} = \begin{pmatrix} 0 & N'' \\ N & 0 \end{pmatrix}.\end{aligned}$$

Therefore we have proved,

THEOREM 4. *If t is even and r^t occurs c times among the elementary factors of Λ, then c is even. Corresponding to each pair of elementary factors r^t, r^t, there is a block matrix $\begin{pmatrix} 0 & N'' \\ N & 0 \end{pmatrix}$ in the canonical form of Λ.*

Theorems 1, 3 and 4 complete the solution of problem (b) when $\epsilon = 1$, $\delta = -1$.

Case 3. The matrices H_{11} and $H_{11}U$ are both skew symmetric so that $H_{11} = U'H_{11}$. If T is the counter unit matrix of order t,

$TU = U'T$ and $H_{11} = Tf(U)$. Since $H_{11} = -H'_{11}$,
$Tf(U) = -f(U)T = -Tf(U)$, so that $f(U) = 0$ and $H_{11} = 0$.

Therefore we need only consider blocks of type (β), $\begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix}\begin{pmatrix} N & 0 \\ 0 & N \end{pmatrix}$

which is of the same nature as (33) and can be reduced to $\begin{pmatrix} 0 & N'' \\ N & 0 \end{pmatrix}$ by (34). Therefore,

THEOREM 5. *If r^t occurs c times among the elementary factors of a skew symmetric pencil Λ, c is even. Corresponding to each pair of elementary factors r^t, r^t there is a block matrix $\begin{pmatrix} 0 & N'' \\ N & 0 \end{pmatrix}$ in the canonical form of Λ.*

3. In sections 1 and 2 we have proved that there exists in all three cases a non-singular matrix P such that

$$P^*\Lambda P = [\Lambda_1, \Lambda_2]$$

where the form of Λ_1 is determined and

$$\Lambda_2 = H(rM + sE), \tag{35}$$

the matrix H being non-singular.

As mentioned in the introduction a normal form for Λ_2 has been determined in case 1 (I) and also in case 2 (II). We now consider case 3. The matrices H and HM are both skew symmetric, so that $HM = M'H$ and, consequently, if Q is any non-singular matrix satisfying the equation

$$QM = M'Q, \tag{36}$$

then $H = QG$, where $GM = MG$.

Let the elementary factors of $rM + sE$ be the homogeneous polynomials

$$p_i(r, s)^{\eta_{ij}}, \qquad (i = 1, 2, \cdots, t;\ j = 1, 2, \cdots, r_i), \tag{37}$$

where $p_i(r, s)$ is irreducible in $K[r, s]$.

We may take M in the canonical form

$$M = [M_1, M_2, \cdots, M_t],$$

where the elementary factors of $rM_k + sE_k$ are the polynomials (37), when $i = k$. Since there exists a non-singular matrix Q_k such that $Q_kM_k = M'_kQ_k$ (I page 478), the matrix $Q = [Q_1, Q_2, \cdots, Q_t]$ is non-singular and satisfies (36). Any matrix G commutative with M is also a diagonal block matrix $[G_1, G_2, \cdots, G_t]$ [11] and consequently

$$H = QG = [Q_1G_1, Q_2G_2, \cdots, Q_tG_t].$$

Therefore it is sufficient to consider a pencil Λ_2 in (35), whose elementary factors are all powers of the same irreducible polynomial $p(r, s)$ i. e. are

---

[11] John Williamson, "The idempotent and nilpotent elements of a matrix," *American Journal of Mathematics*, vol. 58 (1936), pp. 747-758.

$$p(r,s)^{\eta_i}, \qquad \eta_1 \geqq \eta_2 \geqq \cdots \geqq \eta_t.$$

Accordingly we take M in the canonical form

(38) $$M = [L_1, L_2, \cdots, L_t],$$

where

(39) $$L_i = p \cdot E_i + e \cdot U_i, \text{ and write}$$

(40) $$N_i = rL_i + se \cdot E_i.$$

In (39), E_i and U_i are respectively the unit matrix and the auxiliary unit matrix of order η_i; p is the companion matrix of $p(1,\lambda)$; e the unit matrix of the same order as p; and $\cdot$ denotes direct product. (I page 477). For example, if $\eta_i = 3$,

$$L_i = \begin{pmatrix} p & e & 0 \\ 0 & p & e \\ 0 & 0 & p \end{pmatrix}.$$

It has been shown, I page 490, that there exists a non-singular symmetric matrix q such that $qp = p'q$. Hence, if T_i is the counter unit matrix of order η_i,

$$q \cdot T_i L_i = q \cdot T_i (p \cdot E_i + e \cdot U_i) = (p' \cdot E_i + e \cdot U'_i) q \cdot T_i = L'_i q \cdot T_i.$$

Therefore the matrix

$$Q = [q \cdot T_1, q \cdot T_2, \cdots, q \cdot T_t]$$

satisfies (36), and is symmetric.

Let

$$G = (G_{ij}), \qquad (i, j = 1, 2, \cdots, t),$$

be a partition of a matrix G similar to that of M in (38). If G is commutative with M, the form of G is known.[12] In fact, if $\eta_i \geqq \eta_j$,

(41) $$G_{ij} = \begin{pmatrix} S_{ij} \\ 0 \end{pmatrix}, \qquad G_{ji} = (0, S_{ji}),$$

where S_{ij} and S_{ji} are square matrices of order $\eta = \eta_j$, while 0 is the zero matrix of $\eta_i - \eta_j$ rows and η_j columns. Further

(42) $$S_{ij} = \sum_{a=0}^{\eta-1} s_{ija} U_j^a, \qquad S_{ji} = \sum_{a=0}^{\eta-1} s_{jia} U_j^a,$$

where $s_{ijk} = s_{ijk}(p)$ is a polynomial in the matrix p.

Since $H = QG$ is skew symmetric,

$$QG = -G'Q' = -G'Q.$$

[12] John Williamson, *The Idempotent and Nilpotent Elements of a Matrix*, p. 457.

Hence $q \cdot T_i G_{ij} = - G'_{ji} q \cdot T_j$, or, when $i \leqq j$,

$$q \cdot T_i \cdot \begin{pmatrix} S_{ij} \\ 0 \end{pmatrix} = - \begin{pmatrix} 0 \\ S'_{ji} \end{pmatrix} q \cdot T_j.$$

From this we deduce

$$q \cdot T_j S_{ij} = - S'_{ji} q \cdot T_j.$$

But, by direct calculation from (42),

$$q \cdot T_j S_{ji} = S'_{ji} q \cdot T_j,$$

and hence

(43) $$S_{ij} = - S_{ji}.$$

In particular $G_{ii} = S_{ii} = 0$. Let $\eta_1 = \eta_2 = \cdots = \eta_c > \eta_{c+1}$. Since G is non-singular and, when $j > c$, the first column of G_{j1} is zero, G_{i1} must be non-singular for at least one value of i, $2 \leqq i \leqq c$. We may therefore assume that G_{21} is non-singular and, as a consequence of (43), that

$$\begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} = \begin{pmatrix} 0 & G_{12} \\ G_{21} & 0 \end{pmatrix}$$

is non-singular. Therefore by Lemma 6 there exists a non-singular matrix P such that

$$P' \Lambda_2 P = \left[\begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix} \begin{pmatrix} N_1 & 0 \\ 0 & N_1 \end{pmatrix}, H_2 \Lambda_3 \right].$$

Since the block $\begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix} \begin{pmatrix} N_1 & 0 \\ 0 & N_1 \end{pmatrix}$ is of the same type as (33), it may be reduced to the block matrix $\begin{pmatrix} 0 & N_1'' \\ N_1 & 0 \end{pmatrix}$.

We have therefore proved.

THEOREM 6. *Each elementary factor $p(r,s)^\eta$ of a pencil of skew symmetric matrices must occur an even number of times. Corresponding to each pair of elementary factors $p(r,s)^\eta$, $p(r,s)^\eta$, in the canonical form is a matrix block $\begin{pmatrix} 0 & -N' \\ N & 0 \end{pmatrix}$ where N is defined by* (39) *and* (40).

Combining this with the results of sections (1) and (2) we have finally,

THEOREM 7. *Necessary and sufficient conditions, that two pencils of skew symmetric matrices be equivalent under a non-singular congruent transformation in K, are that the two pencils have the same kronecker minimal row indices and the same elementary factors.*

THE JOHNS HOPKINS UNIVERSITY.

SOME REMARKS ON CLASS FIELD THEORY OVER INFINITE FIELDS OF ALGEBRAIC NUMBERS.*

By O. F. G. Schilling.

Mr. M. Moriya recently investigated the theory of finite abelian extensions over infinite fields of algebraic numbers.[1] He has shown that under certain restricting conditions on the infinite algebraic ground field there exists an analog to the classical class field theory: the finite abelian extensions of an infinite field can be characterized by class groups of ideals in the groundfield. His results can be completed in several directions. In this note we shall characterize the finite algebraic number fields by an intrinsic property of the given field: a number field is finite if and only if there exists a finite number of cyclic superfields of some prime degree with a given defining modulus. Furtherfore we shall prove the analog to the theorem on arithmetic progressions for a certain class of infinite fields. Finally we discuss the norm theorem of Hilbert and Hasse. This connects our investigation with A. A. Albert's results on algebras over infinite number fields.[2]

Let k be an infinite field of algebraic numbers over the field P of all rational numbers. The field k can always be approximated by an enumerable tower of finite algebraic number fields k_i over P; that is to say k is the join $\Sigma\, k_i$ of finite fields k_i such that

$$\cdots \subset k_{i-1} \subset k_i \subset \cdots \subset \Sigma\, k_i = k.$$

With the field k there is associated a Steinitz G-number $N(k, \mathrm{P})$, the absolute G-degree of k. The number $N(k, \mathrm{P})$ is defined as the formal least common multiple of all the relative degrees $[h:\mathrm{P}]$ where h is any finite subfield of k. If a prime p divides almost all degrees $[h:\mathrm{P}]$—or almost all degrees $[k_i:k_{i-1}]$ is sufficient too—then we say that p^∞ divides $N(k, \mathrm{P})$. Thus $N(k, \mathrm{P})$ can be uniquely decomposed into an infinite part $N_{\mathrm{inf}}(k, \mathrm{P})$ consisting of the product of all $p^\infty \mid N(k, \mathrm{P})$, and a finite part $N_{\mathrm{fin}}(k, \mathrm{P})$ which consists of the exact powers of those primes p which divide only a finite number of

* Received December 18, 1936.

[1] M. Moriya, "Klassenkörpertheorie für einen unendlichen Zahlkörper." Will appear in the *Journal of the Faculty of Science*, Sapporo (Japan).

[2] A. A. Albert, "Normal division algebras over algebraic number fields not of finite degree," *Bulletin of the American Mathematical Society*, October, 1933.

relative degrees $[h:\mathrm{P}]$. Moriya has shown that exactly all those abelian extensions K of k whose degrees $[K:k]=n$ are prime to $N_{\mathrm{inf}}(k,\mathrm{P})$ are class fields, that is to say the Galois groups $G(K,k)$ of K over k are isomorphic with class groups $\mathfrak{a}/(H(K,k)$ derived from the group $\mathfrak{a}$ of all ideals in k that have inverses.[3]

It is not difficult to construct infinite fields k whose infinite parts $N_{\mathrm{inf}}(k,\mathrm{P})$ of the respective G-degrees $N(k,\mathrm{P})$ are equal to one. Such fields k have then the property that all abelian fields K over k are class fields. There arises the problem of finding properties of these infinite fields k that distinguish them from the finite algebraic number fields.

Now let k_0 be a finite algebraic number field which contain the l-th roots of unity ζ_l $(l\neq 2)$. By $\mathfrak{f}$ we denote an integral ideal of k whose prime divisors we shall later specify, and by ∞ the product over all the infinite prime places of k_0.[4] Assume that $\mathfrak{f}$ is chosen in such a fashion that there exist cyclic extensions Z of k_0 whose degrees are equal to l and whose conductors $\mathfrak{f}(Z,k_0)$ are divisors of $\mathfrak{f}\infty$. Such moduli $\mathfrak{f}\infty$ always exist. The number of different cyclic fields Z of this type shall be denoted by $R(k_0,l,\mathfrak{f}\infty)$.

Let the prime divisors of l in k_0 be $\mathfrak{l}_1,\cdots,\mathfrak{l}_s$; we obtain a decomposition $(l)=\prod_{i=1}^{s}\mathfrak{l}_i^{e(i)}$, and have $\boldsymbol{N}(k_0,\mathrm{P}|\mathfrak{l}_i)=(l)^{f(i)}$ where $\sum_{i=1}^{s}e(i)f(i)=[k_0:P]$. Then the numbers $w(k_0,\mathfrak{l}_i)=e(i)l(l-1)^{-1}$ are always integral.[5] Suppose now that the ideal $\mathfrak{f}=\prod_{(\mathfrak{p},\mathfrak{l})=1}\mathfrak{p}\prod_{\mathfrak{l}_i|l}\mathfrak{l}_i^{v(i)}$ is chosen in such a way that the exponents $v(i)$ are greater than or equal to the numbers $w(k_0,\mathfrak{l}_i)$ and that there occur sufficiently many prime ideals $\mathfrak{p}$. Then there surely exist cyclic fields Z of degree l over k whose conductors $\mathfrak{f}(Z,k)$ divide $\mathfrak{f}\infty$.

LEMMA 1. *If k'_0 is a finite extension of k_0 whose relative degree n' is prime to l then*

$$R(k'_0,l,\mathfrak{f}\infty)>R(k_0,l,\mathfrak{f}\infty).$$

Proof. First we give an explicit expression for the number $R(k_0,l,\mathfrak{f}\infty)$. By $w(\mathfrak{p})$ and $w(\mathfrak{l})$ we denote the number of prime divisors $\mathfrak{p}$ and $\mathfrak{l}$ which divide $\mathfrak{f}$. The number of fundamental basic units of k_0 is equal to $r(k_0)=[k_0:\mathrm{P}]\,2^{-1}-1$ because there exist no real infinite prime places in k_0 and the number of complex infinite prime places is equal to $[k_0:\mathrm{P}]\,2^{-1}$ as the following inclusion shows

---

[3] For details see the Paper of M. Moriya mentioned under no. 1.

[4] For the class field theory over finite number fields see the Tract of H. Hasse in vol. 35 (1930) of the *Jahresberichte der Deutschen Mathematikervereinigung.*

[5] The following formulae can be found in Hasse's Tract, part Ia, § 15.

[6] Cf. Hasse's Tract, part II.

$$\mathrm{P} \subset \mathrm{P}(\zeta_l + \zeta_l^{-1}) \subset \mathrm{P}(\zeta_l) \subseteq k.$$

Now let $\{\alpha\}$ be the multiplicative group of all numbers α in k_0 which are prime to $\mathfrak{f}$. The group $\{\alpha\}$ contains the subgroup $\{\omega\}$ of all numbers ω in k_0 for which $\omega \equiv \alpha^l \pmod{\mathfrak{f}}$ and $(\omega) = \mathfrak{r}^l$ where $\mathfrak{r}$ are non-principal ideals of k_0 which are relatively prime to $\mathfrak{f}$. (Written as $\mathfrak{r} \not\sim 1(k_0)$.) Then the index $[\{\alpha\} : \{\omega\}]$ is equal to the $m(k_0)$-th power of l. Finally $R(k_0, l, \mathfrak{f}\infty)$ becomes equal to $(l^S - 1)(l-1)^{-1}$ where

$$\begin{aligned} S &= w(\mathfrak{p}) + w(\mathfrak{l}) + m(k_0) + \sum_{i=1}^{s} e(i)f(i) - (r(k_0) - 1) \\ &= w(\mathfrak{p}) + w(\mathfrak{l}) + m(k_0) + [k_0 : P]\, 2^{-1}. \end{aligned}$$

Now we wish to calculate the expression S' in k' with respect to the same modulus $\mathfrak{f}\infty$. The prime ideal $\mathfrak{p}$ may be decomposed as $\mathfrak{p} = \prod_{j=1}^{g(..)} \mathfrak{p}'_j{}^{e(j)}$ where $g(..) = g(\mathfrak{p})$; then $w(\mathfrak{p}') \geq w(\mathfrak{p})$ and the exponents $e(j)$ can be normalized to one because they do not matter in the determination of S'. For the prime ideals $\mathfrak{l}_i$ we obtain similarly $\mathfrak{l}_i = \prod_{j=1}^{g(i)} \mathfrak{l}'^{e(i,j)}_{ij}$, and $\Pi\, \mathfrak{l}_i{}^{v(i)}$ goes over into $\prod_{i,j} \mathfrak{l}'^{v(i)e(i,j)}_{ij}$. As a simple calculation shows, we have

$$v(i)e(i,j) \geq w(k'_0, \mathfrak{l}'_{ij}) = e(i,j)w(k_0, \mathfrak{l}_i);$$

hence $w(\mathfrak{l}') \geq w(\mathfrak{l})$.

Moreover, $m(k'_0) \geq m(k_0)$. This is seen as follows. Evidently it is sufficient to show that an ideal $\mathfrak{r} \not\sim 1$ (k_0) cannot become a principal ideal in k'. Assume that $\mathfrak{r} \sim 1$ (k'_0), then $N(k', k|\mathfrak{r}) = \mathfrak{r}^{n'} \sim 1$ (k_0). According to our assumptions, $(n', l) = 1$; hence there exist two integers c and d such that $cn' + dl = 1$. This leads to the relation $\mathfrak{r}^1 = \mathfrak{r}^{cn'+dl} = (\mathfrak{r}^{n'})^c(\mathfrak{r}^l)^d \sim 1$ (k_0), for $\mathfrak{r}^l \sim 1$ (k_0). But this is in contradiction to the assumptions on the ideal $\mathfrak{r}$. Thus we obtain

$$\begin{aligned} S' &= w(\mathfrak{p}') + w(\mathfrak{l}') + m(k'_0) + [k' : \mathrm{P}]\, 2^{-1} \\ &= w(\mathfrak{p}') + w(\mathfrak{l}') + m(k'_0) + n'[k_0 : \mathrm{P}]\, 2^{-1}, \end{aligned}$$

so that certainly $S' > S$ for $n' > 1$. Thus we have

$$R(k'_0, l, \mathfrak{f}\infty) > R(k_0, l, \mathfrak{f}\infty).$$

Now let k be an infinite algebraic extension of k_0 such that $N_{\mathrm{inf}}(k, \mathrm{P})$ is prime to l.

Lemma 2. *The number of cyclic extensions $\mathfrak{Z}$ of degree l over k whose conductors are divisors of $\mathfrak{f}\infty$ is infinite.*

Proof. Each algebraic extension $K = k(\vartheta)$ of k is evidently the join of k and a finite field K_0. For take some finite subfield k_0 of k such that the coefficients belonging to the irreducible equation of ϑ in k lie in it; then $k_0(\vartheta)$ is a field with the asserted property. Thus all cyclic fields Z of degree l are joins of finite cyclic fields Z_j over $k_j \subset k$ with k. Now $R(k_i, l, \mathfrak{f}\infty) > R(k_{i-1}, l, \mathfrak{f}\infty)$ for sufficiently high i, because $([k_i : k_{i-1}], l) = 1$ as a consequence of $(N_{\text{inf}}(k, \mathrm{P}), l) = 1$. Thus there exist infinitely many cyclic fields Z with the described property.

THEOREM 1. *If k is an algebraic number field such that for a prime $l \neq 2$ the class field theory holds, then k is infinite if and only if there exist infinitely many cyclic superfields Z of degree l over k whose conductors are divisors of a given modulus $\mathfrak{f}\infty$ lying in a finite subfield of k.*

Proof. Without loss of generality we can assume that k contains the l-th roots of unity, because $k(\zeta_l)$ is at most of degree $l - 1$ over k. Therefore the assumptions on k are carried over to $k(\zeta_l)$, $N_{\text{inf}}(k, \mathrm{P}) = N_{\text{inf}}(k(\zeta_l), \mathrm{P})$.

Now if k is infinite we denote by k_0 some finite subfield of k such that it contains the l-th roots of unity and that $N(k, k_0)$ is prime to l. The modulus $\mathfrak{f}\infty$ shall be chosen as a modulus of the type investigated before; then from the fact that for l the class field theory holds—that is to say, all cyclic fields of degree l over k can be described uniquely by class groups of ideals in k—it follows that $(N_{\text{inf}}(k, \mathrm{P}), l) = 1$. Hence Lemma 2 can be applied on k, it asserts that there exist infinitely many fields Z whose conductors divide $\mathfrak{f}\infty$. Conversely, if there exist infinitely many fields Z with these properties then k is infinite. This is obvious because a finite field never possesses an infinity of cyclic superfields of degree whose conductors divide a fixed modulus $\mathfrak{f}\infty$.

THEOREM 2. *If K is a class field of degree n over the infinite algebraic number field k then there exists for every divisor f of n, where f is the order of an element of the Galois group $G(K, k)$, an infinity of prime ideals $\mathfrak{p}$ in k which are prime to the discriminant of K over k and whose prime divisors $\mathfrak{P}$ in K have the relative residue class degree f.*

Proof. We chose a finite subfield k_0 in k such that $N(k, k_0)$ is prime to n; this is always possible because K was assumed to be a class field of degree n over k. If $K = k(\vartheta)$ then there exists some finite field k_i in k which contains k_0 such that $K = K_i k$ where $K_i = k_i(\vartheta)$. According to the theorem on arithmetic progressions in k_i there exist infinitely many prime ideals $\mathfrak{p}(i)$ in k which are prime to the discriminant of K_i over k_i and whose divisors $\mathfrak{P}(i, j(i))$ in K_i are of relative residue degrees f over k_i, if f is the order of an arbitrary

element S_i of the Galois group $G(K_i, k_i) \cong G(K, k)$. Now the Artin symbol $(K_i, k_i/\mathfrak{p}(i)) = S_i$. Let $k_{i+\nu}$ be any finite extension of k_i which belongs to the approximation $\{k_i\}$ of k, and let $\mathfrak{p}(i+\nu)$ be any divisor of a fixed $\mathfrak{p}(i)$. The translation theorem of class field theory asserts that $(K_{i+\nu}, k_{i+\nu}/\mathfrak{p}(i+\nu))$ is also of order f because the degree $[k_{i+\nu} : k_i]$ is prime to n. The residue class degree of any divisor $\mathfrak{P}(i+\nu, j(i+\nu))$ is therefore equal to f for any ν. Now take a sequence $\{\mathfrak{p}(i+\nu)\}$ of prime ideals $\mathfrak{p}(i+\nu)$ such that $\mathfrak{p}(i) \subseteq \mathfrak{p}(i+1) \subseteq \cdots \subseteq \mathfrak{p}(i+\nu) \subseteq \cdots$. This sequence determines a prime ideal $\mathfrak{p}$ in k and it is prime to the discriminant of K with respect to k. In the same fashion we determine a chain of prime ideals $\mathfrak{P}(i+\nu, j(i+\nu))$ in $K_{i+\nu}$ such that

$$\text{i)} \quad \mathfrak{P}(i, j(i)) \subseteq \mathfrak{P}(i+1, j(i+1)) \subseteq \cdots \subseteq \mathfrak{P}(i+\nu, j(i+\nu)) \subseteq \cdots$$

$$\text{and ii)} \quad \mathfrak{p}(i+\nu) \subseteq \mathfrak{P}(i+\nu, j(i+\nu)).$$

The limit prime ideal $\mathfrak{P}$ of $\{\mathfrak{P}(i+\nu, j(i+\nu))\}$ in K is then a divisor of $\mathfrak{p}$, and the residue class degree of $\mathfrak{P}$ is equal to f. The equalities

$$f = f(i) = f(i+1) = \cdots f(i+\nu) = \cdots$$

for the respective residue class degrees assert according to Herbrand the norm relation.[7]

Now the number of prime ideals $\mathfrak{p}(i)$ belonging to a fixed order f is infinite; hence the proof of the theorem is complete.

Now let $\mathfrak{p}$ be an arbitrary prime ideal of the infinite algebraic number field k. Then $\mathfrak{p}$ uniquely determines a valuation on the field k. The system of all fundamental sequences $\{\alpha_\mu\}$—α_μ in k—with respect to that valuation form a field $k(\mathfrak{p})$, the so-called derived field of k with respect to $\mathfrak{p}$. If k is equal to $\Sigma\, k_i$, then the intersections $\mathfrak{p}(i) = \mathfrak{p} \cap k_i$ are prime ideals in the finite subfields k_i of k. Form $\Sigma\, k_i(\mathfrak{p}(i))$. This field is in general an infinite algebraic extension of $k_0(\mathfrak{p}(0))$ and it is not closed with respect to $\mathfrak{p}$; but the following lemma holds.

LEMMA 3. *The derived field $k(\mathfrak{p})$ of an infinite algebraic number field $k = \Sigma\, k_i$ is equal to the derived field belonging to the field $\Sigma\, k_i(\mathfrak{p}(i))$ where $k_i(\mathfrak{p}(i))$ denotes the perfect fields of k_i with respect to $\mathfrak{p}(i) = \mathfrak{p} \cap k_i$.*

Proof. According to the construction of the valuation belonging to $\mathfrak{p}$ in k the value groups of $\mathfrak{p}$ in k and of $\mathfrak{p}' = \mathfrak{p} \cap \Sigma\, k_i(\mathfrak{p}(i))$ coincide.[8]

[7] J. Herbrand, "Théorie arithmétique des corps de nombres de degré infini, I. Extensions de degré fini," *Mathematische Annalen*, vol. 106 (1932).

[8] W. Krull, "Idealtheorie in unendlichen Zahlkörpern," *Mathematische Zeitschrift*, vol. 29 (1928).

First we show that $k(\mathfrak{p})$ is contained in the derived field of $\Sigma\, k_i(\mathfrak{p}(i))$. Let $\{\alpha_\mu\}$ be an arbitrary fundamental sequence of elements α_μ in k, it represents an arbitrary element of $k(\mathfrak{p})$. Each of the elements α_μ lies already in a suitable finite subfield of k. Therefore the sequence $\{\alpha_\mu\}$ consists of elements in $\Sigma\, k_i(\mathfrak{p}(i))$; it is also a fundamental sequence with respect to $\mathfrak{p}'$ because the value groups of $\Sigma\, k_i(\mathfrak{p}(i))$ and k resp. $k(\mathfrak{p})$ coincide. Hence $\{\alpha_\mu\}$ lies in the derived field of $\Sigma\, k_i(\mathfrak{p}(i))$.

Conversely, each fundamental sequence $\{\alpha'_\mu\}$ of $\Sigma\, k_i(\mathfrak{p}(i))$ is an element of $k(\mathfrak{p})$. According to the definition of $\Sigma\, k_i(\mathfrak{p}(i))$ each element α'_μ lies already in a finite $\mathfrak{p}(\mu)$-adic field $k_\mu(\mathfrak{p}(\mu)) \supseteq k_\mu$; therefore we always can choose some element α_μ in k which lies arbitrarily near to a'_μ in the sense of the valuation belonging to $\mathfrak{p}'$. The sequence $\{\alpha_\mu\}$ is also a fundamental sequence of $\Sigma\, k_i(\mathfrak{p}(i))$, and according to the definition of the closure of $\Sigma\, k_i(\mathfrak{p}(i))$ we have $\{\alpha_\mu\} = \{\alpha'_\mu\}$. The sequence $\{\alpha_\mu\}$ is a fundamental sequence of k, therefore $\{\alpha'_\mu\} = \{\alpha_\mu\}$ lies in $k(\mathfrak{p})$.

M. Moriya and the author have proved in 2 papers that the class field theories over $\Sigma\, k_i(\mathfrak{p}(i))$ and $k(\mathfrak{p})$ are virtually the same. There exists a one-to-one correspondence between the abelian finite extensions and finite normal algebras over both fields respectively.[9] Therefore it is of no importance in which field arithmetic investigations are made. For convenience we shall work in $\Sigma\, k_i(\mathfrak{p}(i))$ in the following considerations.

THEOREM 3. *If Z is a cyclic class field of degree n over the infinite algebraic number field k, then an algebra $A = (a, Z/k)$ is a complete matric algebra if and only if $A(\mathfrak{p}) = (a, Z/k) \times k(\mathfrak{p})$ are complete matric algebras for all prime divisors $\mathfrak{p}$ of k.*

Proof. According to what we just stated the relation $(a,Z/k) \times k(\mathfrak{p}) \sim k(\mathfrak{p})$ is equivalent to $(a, Z/k) \times \Sigma\, k_i(\mathfrak{p}(i)) \sim \Sigma\, k_i(\mathfrak{p}(i))$. Now let $k_0 \subset k$ be a finite subfield such that $N(k, k_0)$ is prime to n; such a field k_0 always exists if Z is a class field of degree n. Now $Z = k(\vartheta)$; let us take an extension k_* of k_0 such that a and the coefficients belonging to the irreducible equation of ϑ in k lie in k_*, write $Z_* = k_*(\vartheta)$. Hence $(a, Z/k) \sim (a, Z_*/k_*) \times k$.

Now assume that $(a, Z/k) \not\sim k$ although

$$(a, Z/k) \times \Sigma\, k_i(\mathfrak{p}(i)) \sim \Sigma\, k_i(\mathfrak{p}(i))$$

for all $\mathfrak{p}$ in k, the infinite prime spots included. Then also $(a, Z_*/k_*) \not\sim k_*$.

[9] M. Moriya and O. F. G. Schilling, "Zur Klassenkörpertheorie über unendlichen perfekten Körpern," and an additional note. Both to appear in the forthcoming volume of the Sapporo Journal.

According to the fundamental theorem on normal algebras over finite number fields there exists at least one prime *ideal* $\mathfrak{p}_*$ in k_* such that

$$(a, Z_*/k_*) \times k_*(\mathfrak{p}_*) \not\sim k_*(\mathfrak{p}_*).^{10}$$

The field $k_*(\mathfrak{p}_*)$ is a subfield of $\Sigma\, k_i(\mathfrak{p}(i))$. Our assumptions yield

$$(a, Z_*/k_*) \times k_*(\mathfrak{p}_*) \times \Sigma\, k_i(\mathfrak{p}(i)) \sim (a, Z_*/k_*) \times \Sigma\, k_i(\mathfrak{p}(i)) \sim \Sigma\, k_i(\mathfrak{p}(i)),$$

that is to say that $\Sigma\, k_i(\mathfrak{p}(i))$ is a splitting field of $(a, Z_*/k_*) \times k_*(\mathfrak{p}_*)$. The splitting must already be finished in a finite extension $k_\lambda(\mathfrak{p}(\lambda))$ of $k_*(\mathfrak{p}_*)$ because $\Sigma\, k_i(\mathfrak{p}(i))$ is algebraic over $k_*(\mathfrak{p}_*)$. Hence according to the local class field theory over finite $\mathfrak{p}_*$-adic fields, the degree $[k_\lambda(\mathfrak{p}(\lambda)) : k_*(\mathfrak{p}_*)]$ must be a multiple of the exponent belonging to the algebra $(a, Z_*/k_*) \times k_*(\mathfrak{p}_*)$. The latter is a divisor of n and certainly different from one if

$$(a, Z_*/k_*) \times k_*(\mathfrak{p}_*) \not\sim k_*(\mathfrak{p}_*).$$

Hence $([k_\lambda(\mathfrak{p}(\lambda)) : k_*(\mathfrak{p}_*)], n) \neq 1$ and a fortiori $([k_\lambda : k_*], n) \neq 1$. But this contradicts the choice of k_* in k. Therefore we must have

$$(a, Z_*/k_*) \times k_*(\mathfrak{p}_*) \sim k_*(\mathfrak{p}_*)$$

and hence $\mathfrak{p}_*$ cannot be a ramified prime ideal. This means that $(a, Z_*/k_*) \sim k_*$ and a fortiori that $(a, Z/k) \sim (a, Z_*/k_*) \times k \sim k$.—As always the converse is trivial.

Remark. Theorem 3 holds of course for arbitrary simple algebras A over the field k because they all possess cyclic representations.

Now let $k = \Sigma\, k_i$ be an arbitrary infinite algebraic number field. We assume that there exists an algebra A of degree n over k which is not isomorphic with a complete matric algebra over k. The algebra A then can be represented in the form $A_0 \times k$ where A_0 is an algebra A_* over a suitably chosen finite subfield k_* of k. Then we have also $A_* \not\sim k_*$ and therefore there exists at least one prime ideal $\mathfrak{p}(0) = \mathfrak{p}_*$ in the field $k_0 = k_*$ which we shall take as field to start with the approximation $\{k_i\}$ of k, such that

$$A_0 \times k_0(\mathfrak{p}(0)) \not\sim k_0(\mathfrak{p}(0)).$$

The exponent of $A_0 \times k_0(\mathfrak{p}(0))$ may be denoted by $m(\mathfrak{p}_*)$. By $\mathfrak{p}$ we denote any prime divisor of $\mathfrak{p}_*$ in k; then $\mathfrak{p} = \lim \mathfrak{p}(i)$ where

$$\mathfrak{p}_* = \mathfrak{p}(0) \subseteq \mathfrak{p}(1) \subseteq \cdots \subseteq \mathfrak{p}(i) \subseteq \cdots.$$

[10] For the theory of normal algebras see H. Hasse, "Über die Struktur etc., . . . ," *Mathematische Annalen*, vol. 107 (1933).

And the G-degree

$$[\Sigma k_i(\mathfrak{p}(i)) : k(\mathfrak{p}(0))] = N(k, k_0; \mathfrak{p}) = N(k, k_*; \mathfrak{p})$$

is equal to the least common multiple of all the degrees

$$[k_i(\mathfrak{p}(i)) : k_{i-1}(\mathfrak{p}(i-1))].$$

LEMMA 4. *If $A = A_* \times k$ is a proper non-matric algebra of degree n over the infinite field $k = \Sigma k_i$ then there exists at least one prime ideal $\mathfrak{p}$ in k lying over a prime ideal $\mathfrak{p}_*$ of $k_* \subset k$, such that*

$$N(k, k_*; \mathfrak{p}) \not\equiv 0 \pmod{m(\mathfrak{p}_*)}.$$

Proof. The relation $A = A_* \times k \not\sim k$ asserts that k is not a splitting field of A_*. Therefore no finite extension k_i of $k_* = k_0$ is a splitting field of A_*. Hence there exists according to Hasse's criterion on finite splitting fields at least one prime ideal $\mathfrak{p}(0)$ in k_0 such that

$$[k_i(\mathfrak{p}(i)) : k_0(\mathfrak{p}(0))] \not\equiv 0 \pmod{m(\mathfrak{p}(0))}$$

if $\mathfrak{p}(i)$ is any prime divisor of $\mathfrak{p}(0) = \mathfrak{p}_*$ in k_i. If we now select a sequence

$$\mathfrak{p}_* = \mathfrak{p}(0) \subseteq \mathfrak{p}(1) \subseteq \cdots \subseteq \mathfrak{p}(i) \subseteq \cdots$$

then its limit prime ideal $\mathfrak{p}$ has the property that

$$N(k, k_*; \mathfrak{p}) \not\equiv 0 \pmod{m(\mathfrak{p}_*)}.$$

Obviously there exist in general many prime ideals $\mathfrak{p}$ lying over $\mathfrak{p}_*$ for which this relation is fulfilled.

Now we are able to extend Theorem 3 to arbitrary fields k.

THEOREM 4. *Any normal algebra A over k of finite degree is a matric algebra over k if and only if $A \times k(\mathfrak{p}) \sim k(\mathfrak{p})$ for all prime divisors $\mathfrak{p}$ of k.*

Proof. Assume that $A \not\sim k$, in spite of $A \times k(\mathfrak{p}) \not\sim k(\mathfrak{p})$, for all $\mathfrak{p}$. According to Lemma 4 there would exist a finite subfield k_* of k and prime ideals $\mathfrak{p}_*$ and $\mathfrak{p}$ such that $N(k, k_*; \mathfrak{p}) \not\equiv 0 \pmod{m(\mathfrak{p}_*)}$. This contradicts the assumption $A \times k(\mathfrak{p}) \sim k(\mathfrak{p})$ which is equivalent to

$$A \times \Sigma k_i(\mathfrak{p}(i)) \sim \Sigma k_i(\mathfrak{p}(i)),$$

for the latter asserts that

$$[k_j(\mathfrak{p}(j)) : k_*(\mathfrak{p}_*)] \equiv 0 \pmod{m(\mathfrak{p}_*)}$$

for suitable j.

Finally we wish to show by an example that there exist proper division algebras D of degree n over certain infinite algebraic number fields k although the G-degree of k is divisible by n^∞.

Let k_0 be an arbitrary algebraic number field, and let n be an arbitrary positive integer. Suppose that $\mathfrak{p}(0)^\nu$ $(\nu = 1, 2, \cdots, s)$ is any finite set of prime divisors in k containing at least one prime ideal and such that we can attribute to each $\mathfrak{p}(0)^\nu$ a rational fraction of maximal denominator n $\rho(0)^\nu$ mod 1 for which $\sum_\nu \rho(0)^\nu \equiv 0 \pmod 1$, where one of them has exactly the denominator n. Then there exists a uniquely determined division algebra D_0 over k_0 in which the $\mathfrak{p}(0)^\nu$ are ramified and which has exactly the exponent n. We now proceed to construct an infinite algebraic extension k of k_0 such that $D_0 \times k$ is still a division algebra and such that $N(k, k_0) = n^\infty$. According to a theorem of Grunwald there exist infinitely many abelian fields k^i of degree n over k such that the prime divisors $\mathfrak{p}(0)^\nu$ are totally decomposed in each of them.[11] According to the criterion on splitting fields none of the fields k^i is a splitting field of D. The fields

$$k_0,\ k_0k^1 = k_1, = k_1,\ k_1k^2 = k_2, \cdots, k_{i-1}k^i = k_i, \cdots$$

define an infinite field k of G-degree n over k_0. The field k is obviously not a splitting field of D, because no finite subfield of k is splitting field. For the same reason D remains a division algebra. We may mention that

$$k(\mathfrak{p}^\nu) \cong k_0(\mathfrak{p}(0)^\nu)$$

for a prime ideal $\mathfrak{p}^\nu = \lim \mathfrak{p}(i)^\nu$, $\mathfrak{p}(i)^\nu \mid \mathfrak{p}(0)^\nu$ in k_i.

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

[11] W. Grundwald, "Ein allgemeines Existenztheorem für algebraische Zahlkörper," *Journal für Mathematik,* vol. 169 (1933).

ON THE ADDITION OF CONVEX CURVES. II.*

By Richard Kershner.

By the vectorial sum $C_1(+)C_2$ of two convex curves C_1 and C_2 is meant the set of all points which may be represented in at least one way as the vectorial sum of a point on C_1 and a point on C_2. It has been shown by Bohr[1] that $C_1(+)C_2$ is either the closed interior of a convex curve C_E or is the closed annular region between two convex curves C_E and C_I, where C_I lies wholly within C_E. The outer boundary C_E of $C_1(+)C_2$ was discussed by Haviland,[2] who found very precise relationships between C_E and the component curves C_1, C_2 by the use of the Minkowski supporting function. For example, if C_1 and C_2 each possess a continuous positive radius of curvature then so does C_E and, in fact, if $\rho_1(\theta), \rho_2(\theta), \rho_E(\theta)$ are the radii of curvature of C_1, C_2, C_E respectively, at the point of C_1, C_2, C_E where the oriented normal has the inclination θ then $\rho_E(\theta) = \rho_1(\theta) + \rho_2(\theta)$.

Recently the author[3] has investigated the inner boundary curve C_I of $C_1(+)C_2$. The results obtained are similar to those of Haviland but the methods are essentially more complicated. The great distinction between the treatment and the results in the two cases is illustrated by the fact that the curve C_I may possess corners while both C_1 and C_2 are analytic.[4] This fact contrasts remarkably with the above remarks concerning the radius of curvature $\rho_E(\theta)$ of C_E.

The purpose of the present note is a discussion of the possible existence of corners in the curve C_I. Specifically, it will be shown that if C_1 and C_2 are analytic curves, then C_I can have but a finite number of corners. (It will be shown that this number may be arbitrarily large). On the other hand, if it be only required of C_1 and C_2 that they possess radii of curvature which

* Received February 28, 1937.

[1] H. Bohr, "Om Addition of uendelig mange konvekse Kurver," *Danske Videnskabernes Selskab* (Forhandlinger, 1913), pp. 325-366. For a short presentation of the proof of this fact cf. B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function," *Transactions of the American Mathematical Society*, vol. 38 (1935), p. 69.

[2] E. K. Haviland, "On the addition of convex curves in Bohr's theory of Dirichlet series," *American Journal of Mathematics*, vol. 55 (1933), pp. 332-334.

[3] R. Kershner, "On the addition of convex curves," *American Journal of Mathematics*, vol. 58 (1936), pp. 737-746.

[4] Cf., e. g., R. Kershner, "On the values of the Riemann ζ-function on fixed lines $\sigma > 1$," *American Journal of Mathematics*, vol. 59 (1937), pp. 167-174.

can be differentiated infinitely often, then it is possible that C_I have an infinite number of corners. Finally it will be shown that if C_1 and C_2 have each a continuous radius of curvature then the corners of C_I are nowhere dense.

In the sequel it will always be assumed that C_I exists. Then [5] one of the curves C_1, C_2 may be placed in the other, after a rotation through the angle π about the origin, by a translation. It will be assumed that C_1 is the "larger" of the two curves so that C_2 may be placed in C_1, in the manner indicated above. By a point of C_1, C_2, C_I, in the direction θ, or, briefly, a point θ, is meant a point where the oriented normal has the inclination θ. Every point of C_I, except a corner, has a direction in the cases to be considered. A corner of C_I will be said to have the direction (θ_1, θ_2) if θ_1 and θ_2 are respectively the lower and upper limits of the directions of points in the neighborhood of the corner.

Using these notations we prove

LEMMA I. *If C_I has a corner in the direction (θ_1, θ_2) and if C_1 and C_2 have continuous, positive radii of curvature $\rho_1(\theta)$ and $\rho_2(\theta)$ respectively, then $\rho_1(\theta) = \rho_2(\theta + \pi)$ for some θ in $\theta_1 < \theta < \theta_2$; and, on the other hand, if $\rho_1(\theta) \leqq \rho_2(\theta + \pi)$ for some interval $\theta'_1 < \theta < \theta'_2$ then C_I has a corner in the direction (θ_1, θ_2) where $\theta_1 \leqq \theta'_1 < \theta'_2 \leqq \theta_2$.*

Proof. In terms of the mechanical interpretation [6] of C_I it is clear that the existence of a corner of C_I in the direction (θ_1, θ_2) means that the curve C_2, after being rotated through an angle π about the origin, may be placed within C_1 in such a way that it has internal contact with C_1 at the two points of C_1 which have the directions θ_1, θ_2. Lemma I follows [7] immediately from this fact. Lemma I gives immediately

THEOREM I. *If C_1 and C_2 are analytic curves then C_I has at most a finite number of corners.*

For suppose C_I had an infinite number of corners. Then the function $\rho_1(\theta) - \rho_2(\theta + \pi)$, where $\rho_i(\theta)$ is the radius of curvature of C_i, would have zeros clustering as some particular point θ. But since $\rho_i(\theta)$ is regular analytic this would imply $\rho_1(\theta) \equiv \rho_2(\theta + \pi)$ and C_I would not exist.

Now if $\rho(\theta)$ is a positive, continuous, periodic function of θ, with period

[5] R. Kershner, *loc. cit.* 3, Theorem I_0, p. 738.

[6] R. Kershner, *loc. cit.* 3, p. 741.

[7] Cf., e. g., S. Mukhopadhyaya, "Circles incident on an oval of undefined curvature," *Tôhoku Mathematical Journal*, vol. 34 (1931), pp. 115-129.

2π, then there will exist a closed convex curve of which $\rho(\theta)$ is the radius of curvature, if and only if the closure conditions

$$(1)\qquad \int_0^{2\pi} \rho(\theta)\cos\theta d\theta = 0; \qquad \int_0^{2\pi} \rho(\theta)\sin\theta d\theta = 0$$

are satisfied.[8] Using this fact it is very easy to show that the number of corners of C_I may be arbitrarily large even when both C_1 and C_2 are analytic. For let C_2 be a circle of radius r. Let

$$(2)\qquad \rho_1^{(n)}(\theta) = r + 1 + \cos n\theta - \delta_n, \qquad (n = 2, 3, 4, \cdots)$$

where $\delta_n > 0$ is chosen so small that, first of all, $\rho_1^{(n)}(\theta)$ is everywhere positive, so that ((1) being obviously satisfied) there does exist a closed convex curve $C_1^{(n)}$ having $\rho_1^{(n)}(\theta)$ as radius of curvature; and secondly that C_2 may be placed entirely within $C_1^{(n)}$. In order to satisfy the first requirement on δ_n it is obviously enough to choose $\delta_n < r$. To see that the second requirement may be satisfied it is enough to notice that if $\delta_n = 0$ then the corresponding $C_1^{(n)}$ has a radius of curvature never less and sometimes greater than r so that C_2 can be placed entirely within $C_1^{(n)}$ by a known theorem.[9] Now let δ_n be fixed satisfying the above requirements and consider the corresponding $C_1^{(n)}$. There are clearly n disjoint θ-intervals in which the radius of curvature (2) of $C_1^{(n)}$ is less than r. Then, by Lemma I, the inner boundary $C_I^{(n)}$ of the vectorial sum $C_1^{(n)}(+)C_2$ will have a corner in the directions (θ_i, θ'_i) for a set of direction intervals including these n disjoint θ-intervals. In general it is not true that two disjoint intervals in which $\rho_1^{(n)}(\theta) \leqq \rho_2(\theta+\pi)$ correspond to distinct corners but in this case the symmetry of $C_1^{(n)}$ makes it obvious that the n intervals mentioned above actually correspond to n distinct corners. Thus, *the inner boundary $C_I^{(n)}$ of the analytic convex sum $C_1^{(n)}(+)C_2$, where $C_1^{(n)}$ is defined by its radius of curvature* (2) *and C_2 is a circle of radius r, has n corners.*

THEOREM II. *There exist convex curves C_1, C_2 whose radii of curvature $\rho_1(\theta)$, $\rho_2(\theta)$ have infinitely many derivatives and which are such that the vector sum $C_1(+)C_2$ has an inner boundary C_I with infinitely many corners.*

Proof. Let C_2 be again a circle of radius r so that $\rho_2(\theta) \equiv r$. Let $\theta_0, \theta_1, \cdots \theta_n, \cdots$ be an infinite sequence of θ-values such that

$$(3)\qquad \pi/2 = \theta_0 > \theta_1 > \cdots > \theta_n \to 0.$$

Let

$$(4)\qquad \rho_1(\theta) \equiv r \text{ if } \theta_{2n} \geqq \theta \geqq \theta_{2n+1}; \qquad (n = 0, 1, 2, \cdots)$$

[8] W. Blaschke, *Kreis und Kugel*, Leipzig (1916), pp. 115-116.

[9] Cf., e. g., W. Blaschke, *loc. cit.*, p. 116.

and

(5) $$\rho_1(\theta) \equiv r + h_n(\theta) \text{ if } \theta_{2n+1} > \theta > \theta_{2n+2}; \qquad (n = 0, 1, 2, \cdots)$$

where $h_n(\theta) > 0$ is a function[10] (defined only on the interval $\theta_{2n+1} > \theta > \theta_{2n+2}$) for which all derivatives (and the function values) exist, and approach zero as $\theta \to \theta_{2n+1} - 0$ or $\theta \to \theta_{2n+2} + 0$, and such that the first n derivatives are less than $(\theta_{2n+2})^2$ in absolute value. Thus $\rho_1(\theta)$ is defined by (3), (4), (5) for the interval $0 < \theta \leqq \pi/2$. Let the definition of $\rho_1(\theta)$ be completed by the requirement

(6) $$\rho_1(\theta) \text{ is periodic of period } \pi/2.$$

Now $\rho_1(\theta)$ is differentiable infinitely often. This is obvious in the interval $0 < \theta \leqq \pi/2$. Thus, by (6), it is sufficient to show that all derivatives exist at the point $\theta = 0$. But, by (4) for $n = 0$, the left-hand derivatives at $\theta = \pi/2$ and hence, by (6), at $\theta = 0$ are all zero. On the other hand, by (3), (4), (5) and the definition of $h_n(\theta)$, the right-hand derivatives are all zero also. For suppose it has been proved that the k-th derivative at $\theta = 0$ exists and is zero, then the $(k+1)$-th difference quotients are bounded by $(\theta_{2n+2})^2/\theta_{2n+2} \to 0$.

Now, by (6), the closure conditions (1) are obviously satisfied so that there exists a convex curve C_1 of which $\rho_1(\theta)$ is the radius of curvature. The circle C_2 of radius r can be placed entirely within C_1 since $\rho_1(\theta) \geqq r$. Thus the vectorial sum $C_1(+)C_2$ does have an inner boundary curve C_I. But if the "mechanical" interpretation of vectorial addition mentioned above be remembered, it is clear that each of the intervals involved in (4) will be directions of distinct corners of this inner boundary C_I. This completes the proof of Theorem II.

It is noticed that the example was so constructed that the points of C_1 and C_2 which corresponded to a cluster point of corners of C_I were points where the radii of curvature of the two curves were equal. This fact could not be avoided. In fact, it is a direct consequence of Lemma I that if the two curves C_1 and C_2 have continuous radii of curvature $\rho_1(\theta)$ and $\rho_2(\theta)$ and if the inner boundary curve of $C_1(+)C_2$ exists and has infinitely many corners clustering in the direction θ_0 then $\rho_1(\theta_0) = \rho_2(\theta_0 + \pi)$. Thus

THEOREM III. *If C_1 and C_2 have continuous radii of curvature and if the vectorial sum $C_1(+)C_2$ has an inner boundary curve C_I then the corners (if any) of C_I are nowhere dense on the inner curve.*

THE JOHNS HOPKINS UNIVERSITY.

[10] Such a function may be taken in the form

$$h_n(\theta) = k_n \exp[(\theta_{2n+1} - \theta)(\theta_{2n+2} - \theta)]$$

where the constant $k_n > 0$ is chosen sufficiently small.

REAL CANONICAL BINARY TRILINEAR FORMS.*

By Rufus Oldenburger.

1. Introduction. In 1922, E. Schwartz [1] found all of the canonical binary trilinear forms for the class of all non-singular linear transformations in the complex field, and distinguished them by means of algebraic invariants. In 1932, the author found these canonical forms independently, and classified them more briefly according to arithmetic invariants.[2] In the present paper the author obtains all of the canonical binary trilinear forms for the class of all non-singular linear transformations in the field of reals, and a complete invariant system. The number of canonical forms is finite and is one more than the number of such forms for the complex field. The method of treatment depends on the use of arithmetic invariants.

Explicitly, the problem solved in this paper is the following: Given two sets of real constants $\bar{a}_{rst}, a_{rst}$, $r, s, t = 1, 2$, find the conditions on $\bar{a}_{rst}, a_{rst}$, such that there exist real solutions $p_r^\rho, q_s^\sigma, m_t^\tau$ of the equations

$$\bar{a}_{rst} = a_{\rho\sigma\tau} p_r^\rho q_s^\sigma m_t^\tau, \qquad (r, s, t, \rho, \sigma, \tau = 1, 2),$$

for which the determinants $|p_r^\rho|$, $|q_s^\sigma|$, $|m_t^\tau|$ are not zero.

2. Definitions. In another paper the author [3] defined and made a thorough study of ranks of n-way matrices and associated forms. A few of these definitions are given here. The *rank* r_i of a 3-way matrix $\boldsymbol{A} = (a_{ijk})$, $i, j, k = 1, 2$, and its associated trilinear form $F = a_{ijk}x_iy_jz_k$; $i, j, k = 1, 2$; is the rank of the 2-way matrix

$$\begin{pmatrix} a_{111} & a_{112} & a_{121} & a_{122} \\ a_{211} & a_{212} & a_{221} & a_{222} \end{pmatrix}.$$

The ranks r_j, r_k are defined similarly. Assume that a_{ijk} $(i, j, k = 1, 2)$ are not

* Received June 16, 1936; revised November 11, 1936.

[1] E. Schwartz, "Über binäre trilineare Formen," *Mathematische Zeitschrift*, vol. 12 (1922), pp. 18-35.

[2] R. Oldenburger, "On canonical binary trilinear forms," *Bulletin of the American Mathematical Society*, vol. 38 (1932), pp. 385-387. In this paper a complete bibliography of earlier papers on binary trilinear forms is given.

[3] R. Oldenburger, "Composition and rank of n-way matrices and multilinear forms," *Annals of Mathematics*, vol. 35 (1934), pp. 622-657.

all zero. The 3-way *rank* $r[jk, i]$ of $\boldsymbol{A}$ and F is defined to be 2 or 1 according as the quantities.[4]

$$(1) \qquad \begin{vmatrix} a_{111} & a_{112} \\ a_{121} & a_{122} \end{vmatrix}, \qquad \begin{vmatrix} a_{211} & a_{212} \\ a_{221} & a_{222} \end{vmatrix}, \qquad \begin{vmatrix} a_{111} & a_{212} \\ a_{121} & a_{222} \end{vmatrix} + \begin{vmatrix} a_{211} & a_{112} \\ a_{221} & a_{122} \end{vmatrix}.$$

are not all zero or are all zero. The ranks $r[ij, k]$, $r[ik, j]$ are defined similarly. Evidently $r[jk, i] = r[kj, i]$. These ranks are *invariant* under non-singular linear transformations on F.

In this paper two binary trilinear forms $F = a_{ijk}x_iy_jz_k$, $G = b_{pqr}x'_py'_qz'_r$ and their associated matrices will be said to be *equivalent* if there exist transformations

$$x_i = a_{ip}x'_p, \qquad y_j = b_{jq}y'_q, \qquad z_k = c_{kr}z'_r,$$

where the square matrices (a_{ip}), (b_{jq}), (c_{kr}) of the second order are non-singular, and these transformations bring F into G. Similarly if F, G are bilinear forms.

3. The canonical forms for which $r_i = r_j = r_k = 2$. By a theorem of another paper[5] if one of the ranks $r[ij, k]$, $r[jk, i]$, $r[ik, j]$ of $\boldsymbol{A}$ is 1, at least one of the ranks r_i, r_j, r_k of $\boldsymbol{A}$ is 1. Hence

$$r[ij, k] = r[jk, i] = r[ik, j] = 2.$$

Let $\boldsymbol{A}_1 = (a_{1jk})$, $\boldsymbol{A}_2 = (a_{2jk})$, $j, k = 1, 2$. Since the coefficients of ρ^2, $\rho\sigma$, σ^2 in the determinant $| \rho\boldsymbol{A}_1 + \sigma\boldsymbol{A}_2 |$ are the quantities (1), it follows that

$$| \rho\boldsymbol{A}_1 + \sigma\boldsymbol{A}_2 | \not\equiv 0.$$

If $\boldsymbol{A}_2$ is non-singular, while $\boldsymbol{A}_1$ is singular make the transformation $x_1 = x'_2$, $x_2 = x'_1$ on the form F. If $\boldsymbol{A}_1, \boldsymbol{A}_2$ are both singular and the matrix $(\rho\boldsymbol{A}_1 + \sigma\boldsymbol{A}_2)$ is non-singular, then $\rho\sigma \neq 0$. Let the bilinear forms $a_{1jk}y_jz_k$, $a_{2jk}y_jz_k$, $j, k = 1, 2$, be denoted by F_1, F_2 respectively. Let

$$F' = a_{ijk}x'_iy'_jz'_k = x'_1F'_1 + x'_2F'_2$$

denote a form for which

$$F'_1 = \rho F_1 + \sigma F_2, \qquad F'_2 = F_2,$$

where ρ, σ are chosen so that $|\rho\boldsymbol{A}_1 + \sigma\boldsymbol{A}_2| \neq 0$. Equating F' to $F = x_1F_1 + x_2F_2$, we obtain

[4] These are 3-way determinants of the second order.

[5] R. Oldenburger, *Annals of Mathematics*, vol. 35 (1934), p. 649.

$$(2) \qquad \begin{aligned} x_1 &= \rho x'_1, \\ x_2 &= \sigma x'_1 + x'_2. \end{aligned}$$

The non-singular transformation (2), therefore, reduces F to a form F' for which (a'_{1jk}) is non-singular. Since in every case F is equivalent to a form $F' = x'_1F'_1 + x'_2F'_2$, where F'_1 is non-singular, it is no restriction to assume in what follows that F_1 is non-singular.

The pair of bilinear forms F_1, F_2 is now equivalent in the field of reals to the canonical pair [6]

$$(3) \qquad F_1 = y_1z_1 + y_2z_2, \qquad F_2 = y_1z_2 + ay_2z_1 + by_2z_2.$$

It is to be noted that any pair of binary bilinear forms is rationally equivalent, in the non-singular case, to the pair (3) or to the pair

$$(4) \qquad y_1z_1 + y_2z_2, \qquad a(y_1z_1 + y_2z_2).$$

Now F_1, F_2 are not equivalent to (4), since, then, the form $F = x_1F_1 + x_2F_2$ has $r_i = 1$.

In what follows in this section, we shall assume that $F = x_1F_1 + x_2F_2$, where F_1, F_2 are as given in (3). We shall consider three cases.

Case 1. $b^2 + 4a > 0$. In the field of reals, the determinant $D = | \rho A_1 + \sigma A_2 |$ factors into distinct linear factors $(\alpha\rho + \beta\sigma)$, $(\gamma\rho + \delta\sigma)$. Let

$$(5) \qquad \rho' = \alpha\rho + \beta\sigma, \qquad \sigma' = \gamma\rho + \delta\sigma.$$

Then $D = \rho'\sigma'$. By another paper of the author [7] the transformation (5) corresponds to a non-singular linear transformation on the x's of F, giving a new form $F' = a'_{ijk}x'_iy_jz_k$, for which

$$(6) \qquad | \rho'a'_{1jk} + \sigma'a'_{2jk} | = \rho'\sigma'.$$

Since the coefficients of ρ'^2 and σ'^2 vanish in (6), the 2-way matrices (a'_{1jk}), (a'_{2jk}) are singular. The form $F'_1 = a'_{1jk}y_jz_k$ is evidently equivalent to a form

$$F_1'' = y_1''z_1''.$$

Simultaneously $F_2' = a'_{2jk}y_jz_k$ transforms into a form

$$F_2'' = ey_1''z_1'' + fy_1''z_2'' + gy_2''z_1'' + hy_2''z_2''.$$

---

[6] L. E. Dickson, *Modern Algebraic Theories*, pp. 89-97.

[7] R. Oldenburger, *Transactions of the American Mathematical Society*, vol. 39 (1936), pp. 432-433.

Hence F' is equivalent to

$$F'' = x_1''F_1'' + x_2''F_2''.$$

Since F_2'' is singular, we can write

$$F_2'' = (\alpha y_1'' + \beta y_2'')(\gamma z_1'' + \delta z_2'').$$

If $\beta, \delta \neq 0$, making the non-singular transformations

$$\begin{aligned} y_1''' &= y_1'', & y_2''' &= \alpha y_1'' + \beta y_2'', \\ z_1''' &= z_1'', & z_2''' &= \gamma z_1'' + \delta z_2'' \end{aligned}$$

on F'', we obtain the canonical form

$$R = x_1 y_1 z_1 + x_2 y_2 z_2,$$

where, for simplicity, the primes on the variables have been removed.

If $\beta = 0$, we can write

$$F'' = y_1''B,$$

where B is a bilinear form in the x's and z's; whence the rank r_j of F'' is 1. Similarly if $\delta = 0$

$$F'' = z_1''Q,$$

where Q is bilinear in the x's and y's and r_k of F'' is 1. In either case we obtain a contradiction of the assumption $r_j = r_k = 2$.

Case 2. $b^2 + 4a = 0$. In this case, $|\rho A_1 + \sigma A_2|$ is a perfect square. The form $F = x_1F_1 + x_2F_2$ defined by (3) is now

$$F = x_1(y_1z_1 + y_2z_2) + x_2(y_1z_2 - g^2y_2z_1 + 2gy_2z_2),$$

where $g = b/2$. Assume that $b \neq 0$. Making the non-singular transformations

$$\begin{aligned} x_1 &= g(x'_2 - x'_1), & x_2 &= x'_1, \\ y_1 &= y'_2, & y_2 &= -(y'_1 + y'_2)/g, \\ z_1 &= (z'_1 - z'_2)/g, & z_2 &= -z'_2 \end{aligned}$$

on F, we obtain

$$L = x_1y_1z_1 + x_2y_1z_2 + x_2y_2z_1,$$

where we have dropped the primes in L.

If $b = 0$, then $a = 0$. Interchanging x_1, x_2 and z_1, z_2 in F, we obtain L.

Case 3. $b^2 + 4a < 0$. In this case, $|\rho A_1 + \sigma A_2|$ does not factor in the field of reals. Let

$$M = x'_1F'_1 + x'_2F'_2,$$

where

(7) $$F'_1 = y'_1z'_1 + y'_2z'_2, \qquad F'_2 = y'_1z'_2 - y'_2z'_1.$$

We shall prove that $F = x_1F_1 + x_2F_2$, defined by (3), is equivalent to M. Apply to M the non-singular linear transformation

(8) $$x'_1 = \rho x_1 + \tau x_2, \qquad x'_2 = \sigma x_1 + \xi x_2.$$

This gives the new form

$$M' = x_1(\rho F'_1 + \sigma F'_2) + x_2(\tau F'_1 + \xi F'_2).$$

For F to be equivalent to M in the field of reals, it is evidently necessary and sufficient that there exist real quantities ρ, σ, τ, ξ, such that, if we write

$$\Delta = \begin{vmatrix} \rho & \sigma \\ \tau & \xi \end{vmatrix},$$

then

(9) $$\Delta \neq 0,$$

and there exist real non-singular transformations on the y's and z's so that M' becomes F. Then there must exist real values of ρ, σ, τ, ξ satisfying (9) such that the pair of bilinear forms $\rho F'_1 + \sigma F'_2$, $\tau F'_1 + \xi F'_2$ is equivalent under non-singular transformations on the y's and z's in the field of reals to F_1, F_2. This equivalence is satisfied *if and only if these pairs of forms have the same invariant factors.*[8] Let

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ a & b \end{pmatrix}.$$

The characteristic matrix of F_1, F_2 is

$$(\lambda \boldsymbol{I} + \mu \boldsymbol{B}) = \begin{pmatrix} \lambda & \mu \\ a\mu & \lambda + b\mu \end{pmatrix},$$

which has the unique invariant factor

(10) $$|\lambda \boldsymbol{I} + \mu \boldsymbol{B}| = \lambda^2 + b\lambda\mu - a\mu^2.$$

The characteristic determinant of $\rho F'_1 + \sigma F'_2$, $\tau F'_1 + \xi F'_2$ is

(11) $$|\lambda(\rho \boldsymbol{I} + \sigma \boldsymbol{A}) + \mu(\tau \boldsymbol{I} + \xi \boldsymbol{A})| = \lambda^2 Q(\rho, \sigma) + 2\lambda\mu B(\rho, \sigma, \tau, \xi) + \mu^2 Q(\tau, \xi),$$

where $\quad Q(\rho, \sigma) = \rho^2 + \sigma^2, \qquad B(\rho, \sigma, \tau, \xi) = (\rho\tau + \sigma\xi).$

[8] L. E. Dickson, *Modern Algebraic Theories*, p. 115.

The invariant factor (10) is equal to the corresponding invariant factor of $\rho F'_1 + \sigma F'_2, \tau F'_1 + \xi F'_2$ if and only if the coefficients of (11) are proportional to those of (10). Then there must exist real values of ρ, σ, τ, ξ satisfying (9), and a real $k \neq 0$ such that

$$(12) \qquad Q(\rho,\sigma) = k, \qquad Q(\tau,\xi) = -ka, \qquad B(\rho,\sigma,\tau,\xi) = kb/2.$$

Since $b^2 + 4a < 0$, we have $a < 0$. It follows, since $Q(\rho,\sigma)$ is a positive definite quadratic form, that $Q(\rho,\sigma)$ represents $k, -ka$ where $k > 0$. If $\rho_1, \sigma_1, \tau_1, \xi_1$ are a set of real values satisfying (12) for a given real $k \neq 0$, the real quantities

$$\rho = \frac{\rho_1}{\sqrt{k}}, \qquad \sigma = \frac{\sigma_1}{\sqrt{k}}, \qquad \tau = \frac{\tau_1}{\sqrt{k}}, \qquad \xi = \frac{\xi_1}{\sqrt{k}}$$

satisfy (12), where in (12) we set $k = 1$. We therefore restrict our study of solutions of (12) to the case $k = 1$. Solving (12_1), (12_2) with $k = 1$ we obtain

$$(13) \qquad \rho = \pm\sqrt{1-\sigma^2}, \qquad \tau = \pm\sqrt{-a-\xi^2}.$$

Substituting the solutions (13) in (12_3) with $k = 1$, we obtain

$$(14) \qquad \pm\sqrt{(1-\sigma^2)(-a-\xi^2)} = (b - 2\sigma\xi)/2.$$

Set

$$(15) \qquad \xi = k_1\sqrt{-a}.$$

Since $a < 0$, ξ is real if k_1 is real. Assume henceforth that k_1 is real. Substituting (15) in (14), we obtain the following solution

$$(16) \qquad \sigma = \frac{-bk_1 \pm \sqrt{(b^2+4a)(k_1^2-1)}}{-2\sqrt{-a}},$$

and from (13) the solution

$$(17) \qquad \tau = \pm\sqrt{a(k_1^2-1)}.$$

Since $b^2 + 4a$, $a < 0$, the quantities σ, τ are real if and only if

$$(18) \qquad k_1^2 \leqq 1.$$

Substituting for ρ, ξ, τ from (13_1), (15), (17) in Δ as given above (9), we find that $\Delta \neq 0$ if

$$\sigma^2 \neq k_1^2.$$

Substituting (16) in the relation

$$\sigma = \pm k_1,$$

transposing terms, squaring, and simplifying, we obtain

$$k_1^2(4\sqrt{-a} \pm b) = -\frac{(b^2+4a)}{2\sqrt{-a}}. \tag{19}$$

Since the right member of (19) is not zero, if there exist solutions of (19), the left member is also $\neq 0$ and, for a given value of the $\pm$ sign, (19) is of the form

$$\alpha k_1^2 = \beta, \qquad \alpha, \beta \neq 0,$$

which has at most two real solutions for k_1.

By (13_1), ρ is real if and only if $\sigma^2 \leqq 1$; whence, by (16), assuming that $k_1^2 \leqq 1$, so that σ is real,

$$-1 \leqq \frac{bk_1 \pm \sqrt{(b^2+4a)(k_1^2-1)}}{2\sqrt{-a}} \leqq 1. \tag{20}$$

Taking the value of the $\pm$ sign in (20) to be $+$, the right inequality of (20) can be reduced to

$$\sqrt{(b^2+4a)(k_1^2-1)} \leqq 2\sqrt{-a} - bk_1. \tag{21}$$

If $b > 0$, the right member of (21) is $\geqq 0$ for

$$\frac{2\sqrt{-a}}{b} \geqq k_1, \tag{22}$$

and, if $b < 0$, that member is $\geqq 0$ for

$$\frac{2\sqrt{-a}}{b} \leqq k_1. \tag{23}$$

If the right member of (21) is $\geqq 0$, and $k_1^2 \leqq 1$, we can square both sides of (21). Simplifying the resulting inequality, we obtain

$$(b - 2\sqrt{-a}\ k_1)^2 \geqq 0,$$

which is satisfied for every real value of k_1.

If $b = 0$, $\sigma = \pm\sqrt{1-k_1^2}$, and $\rho = \pm k_1$, whence ρ is real for every real value of k_1.

Evidently, there is an unlimited number of real values of k_1 satisfying (18), (22) or (23), and not satisfying (19). Also, for any solutions of σ, ξ from (15), (16), the $\pm$ signs in ρ, τ can always be chosen so that (14) is satisfied. We have now proved that in every case we can chose k_1 so that

ρ, σ, τ, ξ are real, $\Delta \neq 0$, and (12) is satisfied. *Hence F is equivalent to M for all a, b such that $b^2 + 4a < 0$.*

4. The canonical forms for which $r_i = 1$. Assume that $r_i = 1$, $r_j = r_k = 2$. We can reduce the form $F = a_{ijk}x_iy_jz_k$, $i, j, k = 1, 2$, at once to x_1B where B is bilinear in y and z and of rank 2. Reducing B to canonical form we obtain

$$H = x_1y_1z_1 + x_1y_2z_2.$$

Assume that $r_i = r_j = r_k = 1$. F can be reduced at once to

$$K = x_1y_1z_1.$$

No form with $r_i = r_j = 1$, $r_k = 2$ exists. We have therefore treated all cases.

5. Fundamental theorems of equivalence. We have proved

THEOREM 1. *Two binary trilinear forms $F = a_{ijk}x_iy_jz_k$ and $G = b_{ijk}x'_iy'_jz'_k$ are equivalent in the field of reals, if and only if they have the same ranks* r_i, r_j, and r_k, and, if $r_i = r_j = r_k = 2$, *the determinants* $|\rho a_{1jk} + \sigma a_{2jk}|$ *and* $|\rho b_{1jk} + \sigma b_{2jk}|$ *have both*

(a) *distinct real linear factors,*

or

(b) *coincident real linear factors,*

or

(c) *no real linear factors.*

THEOREM 2. *In the field of reals, a binary trilinear form $F = a_{ijk}x_iy_jz_k$ is equivalent to one of the following canonical forms:*

(a) $R = x_1y_1z_1 + x_2y_2z_2$, *if* $r_i = r_j = r_k = 2$, *and part* (a) *of Theorem* 1 *is satisfied;*

(b) $L = x_1y_1z_1 + x_2y_1z_2 + x_2y_2z_1$, *if* $r_i = r_j = r_k = 2$, *and part* (b) *of Theorem* 1 *is satisfied;*

(c) $M = x_1y_1z_1 + x_1y_2z_2 + x_2y_1z_2 - x_2y_2z_1$, *if* $r_i = r_j = r_k = 2$, *and part* (c) *of Theorem* 1 *is satisfied;*

(d) $H = x_1y_1z_1 + x_1y_2z_2$, *if* $r_i = 1$, $r_j = r_k = 2$.

(e) $K = x_1y_1z_1$, *if* $r_i = r_j = r_k = 1$.

6. Note concerning M. In the theory of forms, an arithmetic invariant called "factorization rank" plays an important rôle. The factorization ranks

of R, L, H, K have been studied elsewhere[9] by the author. The factorization rank of M is 3, since the matrix (m_{ijk}) of M can be written in the form

$$(m_{ijk}) = (\sum_{\alpha=1}^{3} a_{\alpha i} b_{\alpha j} c_{\alpha k}), \tag{24}$$

where

$$(a_{\alpha i}) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad (b_{\alpha j}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad (c_{\alpha k}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ -1 & 1 \end{pmatrix},$$

and not in the form (24), where the range of α is 1, 2.

7. *Reductions.* The transformations *reducing any trilinear form to canonical form for the field of reals can be written down at once* from the theory of this paper and known theory of bilinear forms.

ARMOUR INSTITUTE OF TECHNOLOGY.

---

[9] R. Oldenburger, "On arithmetic invariants of binary cubic and binary trilinear forms," *Bulletin of the American Mathematical Society*, vol. 42 (1936), pp. 871-873.

A REMARK ON A THEOREM OF ARZELÀ.*

By Philip Hartman.

Let I denote a bounded interval, and $\{f_n(x)\}$ a sequence of functions defined on I such that (i) $\{f_n(x)\}$ is uniformly bounded on I and (ii) every $f_n(x)$ is continuous on I. Condition (i) implies that for every enumerable subset C of I there exists a subsequence of $\{f_n(x)\}$ which is convergent at every point of C. Since C may be chosen dense on I, it follows from a standard theorem of Arzelà that if (i) is satisfied and (ii) is replaced by the more stringent condition that $\{f_n(x)\}$ be equicontinuous on I, then $\{f_n(x)\}$ contains a subsequence which is uniformly convergent on I. The question now arises whether or not (i) and (ii) alone imply the existence of a subsequence which is convergent on I. This question will be answered in the negative by proving a sharper statement to the effect that *there exist sequences* $\{f_n(x)\}$ *which satisfy* (i) *and* (ii) *but are such that every subsequence of* $\{f_n(x)\}$ *is divergent almost everywhere.* For instance, *every* subsequence of the sequence

$$\sin x, \sin 2x, \cdots, \sin nx, \cdots$$

will be shown to be divergent almost everywhere.

Let $\{k_n\}$ be any increasing sequence of positive integers. A theorem of Hardy and Littlewood (*Acta Mathematica,* vol. 37 (1914), p. 181) states that there exists a set $S = S(\{k_n\})$ of measure 1 in $[0, 1]$ such that if θ is a point of S, then the sequence of numbers $\{(k_n\theta)\}$, where $(k_n\theta)$ denotes the fractional part of $k_n\theta$, is dense in $[0, 1]$. It follows that the sequence $\{\sin 2\pi k_n x\}$ is divergent at each point of S; for if θ is a point of S, the sequence of numbers $\{\sin 2\pi k_n\theta\}$ is dense in $[-1, 1]$.

Obviously, the above remarks also apply if $\sin x$ is replaced by any other non-constant, continuous, periodic function $f(x)$ and $f_n(x)$ is defined to be $f(nx)$.

The Johns Hopkins University.

* Received January 6, 1937.

ons
ery
ble
ery
ard
ore
ins
ses
ich
ng
Ech
ent

of
st
nt
al
is
rs

er
to